Author:
Cano Ramirez, Dora

Title:
Integration of circadian and environmental information during signalling to chloroplasts

INTEGRATION OF CIRCADIAN AND ENVIRONMENTAL INFORMATION DURING SIGNALLING TO CHLOROPLASTS

DORA LUZ CANO RAMIREZ

A DISSERTATION SUBMITTED TO THE UNIVERSITY OF BRISTOL IN ACCORDANCE WITH THE REQUIREMENTS FOR AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN THE FACULTY OF SCIENCE

SCHOOL OF BIOLOGICAL SCIENCES
UNIVERSITY OF BRISTOL
UNITED KINGDOM

WORD COUNT = CIRCA 35, 000

ABSTRACT

The coordination of biological processes with daily and seasonal changes in the environment is important for the survival of photosynthetic organisms such as plants. To achieve this, temporal information must be extracted from the external environment and integrated into the biological system for correct co-ordination of gene expression.

A signalling pathway was previously identified that communicates temporal and light quality information from the circadian oscillator to the chloroplast in *Arabidopsis thaliana*. This signalling mechanism operates through the action of a nuclear encoded protein called SIGMA FACTOR5 (SIG5), which confers promoter specificity to the plastid encoded plastid RNA polymerase (PEP) in the chloroplast where it causes rhythmic transcription of *psbD*. In Arabidopsis, there are six sigma factors encoded in the nucleus (SIGMA FACTOR1 (SIG1)-SIG6). I conducted structural homology modelling of the six sigma factors to gain insights about their specificity and regulation.

I also identified in the laboratory a novel low-temperature signalling pathway involving SIG5 that underpins optimum plant performance under both low and freezing temperatures. I demonstrated that this pathway increases freezing tolerance and photosynthetic efficiency at low temperatures, identified upstream and downstream regulators of the pathway, and demonstrated close integration with the circadian oscillator. This suggests the pathway is of considerable biological importance in plants and could be targeted to improve crop yields under uncertain climate conditions.

It is in this context that I then investigated if the pathway operates under fluctuating natural conditions (e.g. during the day and across seasons) using *Arabidopsis halleri* subsp. *gemmifera,* which is a perennial relative of the annual *A. thaliana*. The data revealed seasonal differences in the amplitude of gene transcript accumulation in the signalling pathway and further analysis through statistical modelling revealed the major environmental inputs to the pathway under natural fluctuating conditions.

Finally, using the rare algae *Aegagropila linnaei* (marimo) daily rhythms of buoyancy driven by photosynthesis were found to be a novel output of the circadian oscillator in plants.

DEDICATION AND ACKNOWLEDGEMENTS

My supervisor Dr. Antony Dodd, an incredible scientist and person that believed in me since the beginning. His excellent guidance has been invaluable, and I will always be grateful for all his support.

Prof. Hiroshi Kudoh that has been the kindest person I have come across and an excellent scientist. I would like to thank the lovely people at the Centre for Ecological Research: Dr. Haruki Nishio and Dr. Jiro Sugisawa for their help at collecting samples in the field. Dr. Mie Honjo for her support and shipping the samples from Japan. Dr. Tomoaki Muranaka for his time and patience while conducting the statistical modelling.

Noriane Simon for her friendship, supporting me through the toughest times of my journey and collecting samples in the field.

Heather and Marc Knight for their support, hosting me in their lab and the opportunity to do science with people I admired from years ago. Paige E. Panter in their laboratory for her patience at teaching me electrolyte leaking assays.

Prof. Marina Gavilanes who always supported me through thick and thin and always welcomes me back into her lab.

My main funding body Consejo Nacional de Ciencia y Tecnología (CONACYT, México) and Secretaría de Educación Pública (SEP) - Dirección General de Relaciones Internacionales (DGRI).

Also, funding provided by the Daiwa Anglo-Japanese Foundation, the Biotechnology and Biological Sciences Research Council, the Bristol Centre for Agricultural Innovation, Alumni Foundation from the University of Bristol, the Society for Experimental Biology, the Company of Biologists and the International Hardship Fund from the University of Bristol.

Al honorable pueblo mexicano.

To squishy :3 gudetama, laika, family and friends

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: [redacted] DATE: 22 November 2018

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF ABBREVIATIONS

1 INTRODUCTION 1

1.1 The Arabidopsis circadian clock 1

1.2 Integration of circadian and environmental information 2

1.2.1 Light 2

1.2.2 Temperature and circadian regulation 4

1.3 Chloroplast genome 6

1.4 Chloroplast gene expression 8

1.5 Retrograde signalling 9

1.6 Low temperature responses and adaption to freezing in the chloroplast 12

1.7 Molecular biology in plants under natural environments 13

1.8 Photosynthesis and circadian clocks in the green lineage 14

2 MATERIALS AND METHODS 17

2.1 CHAPTER 3: THREE-DIMENSIONAL STRUCTURAL MODELLING OF ARABIDOPSIS SIGMA FACTORS 17

2.1.1 Multiple sequence alignment 17

2.1.2 Homology modelling 18

2.1.3 pKa prediction 18

2.1.4 Macromolecular electrostatics determination 19

2.1.5 DNA binding sites 19

2.1.6 Phosphorylation site prediction 19

2.2 CHAPTER 4: SIG5 COMMUNICATES INFORMATION TO THE CHLOROPLAST GENOME CONCERNING LOW TEMPERATURE CONDITIONS 20

2.2.1 Plant material and growth conditions 20

2.2.2 Electrolyte leakage assays 21

2.2.3 RNA extraction and qRT-PCR 21

2.2.4 Measurement of PSII photosynthetic efficiency 22

2.2.5 Lipid composition analysis 22

2.3 CHAPTER 5: SIG5-MEDIATED SIGNALLING TO CHLOROPLASTS UNDER NATURAL CONDITIONS 23

2.3.1 Plant material 23

2.3.2 Luciferase transient expression assay 23

2.3.3 Bioluminescence imaging of promoter-luciferase reporter 23

2.3.4 Measurement of light spectrum and temperature under field conditions .24

2.3.5 Field study time courses 24

2.3.6 Preparation of RNA and qRT-PCR from *A. halleri* samples 25
2.4 CHAPTER 6: TO FLOAT OR NOT TO FLOAT 25
2.4.1 Plant material and growth conditions 25
2.4.2 Microscopy 26
2.4.3 Buoyancy assay 26
2.4.4 Measurement of PSII photosynthetic efficiency 26
2.4.5 Delayed chlorophyll fluorescence analysis 27
2.4.6 Quantitative timecourse analysis 28
3 Three-dimensional structural modelling of Arabidopsis sigma factors 29
3.1 Background 29
3.2 Results 31
3.2.1 Three-dimensional homology modelling 33
3.2.2 Prediction of sigma factor phosphorylation sites 45
3.3 Discussion 48
3.3.1 Regulation of plant sigma factors by phosphorylation 49
4 SIG5 COMMUNICATES INFORMATION TO THE CHLOROPLAST GENOME CONCERNING LOW TEMPERATURE CONDITIONS 52
4.1 Background 52
4.2 RESULTS 53
4.2.1 *SIG5* was induced strongly in response to low temperature 53
4.2.2 SIG5 communicates information to the chloroplast genome concerning low temperature and this is gated by the circadian oscillator 55
4.2.3 HY5/HYH are upstream of the pathway 58
4.2.4 SIG5 does not regulate the CBF pathway 61
4.2.5 *SFR2* is rhythmic, responds to cold and is SIG5-dependent 63
4.2.6 A novel SIG5-dependent retrograde signal 64
4.2.7 SIG5 is important to maintain photosynthetic performance (PSII) under low temperature. 67
4.2.8 Initial damage by freezing assayed through electrolyte leakage 69
4.2.9 SIG5 is important for freezing tolerance 69
4.2.10 Lipid remodelling in chloroplasts as an output of the pathway. 71
4.3 DISCUSSION 74
4.3.1 On the identity of the retrograde signal 74
4.3.2 SIG5 in the regulation of eukaryotic lipid biosynthesis. 78
5 SIG5-MEDIATED SIGNALLING TO CHLOROPLASTS UNDER NATURAL CONDITIONS 79
5.1 Background 79
5.2 RESULTS 81

5.2.1 *SIG5* promoter has similar activity in *A. thaliana* and *A. halleri* 81
5.2.2 SIG5-mediated signalling to chloroplasts under natural conditions 82
5.2.3 Local manipulation of environmental conditions in the field 86
5.2.4 Signalling pathway dynamic model 90
5.2.5 Time delay 92
5.2.6 Environmental data 95
5.2.7 Model evaluation adding the environmental input terms 97
5.2.8 Major determinants of the signalling pathway 101
5.2.9 The role of R:FR in the local environmental manipulation experiments 106
5.2.10 A consensus dynamic model for SIG5-mediated environmental signal integration 109
5.3 Conclusions 110
5.4 Discussion 110
6 TO FLOAT OR NOT TO FLOAT 113
6.1 Background 113
6.2 Results 114
6.3 Discussion 122
6.3.1 Further experiments 123
6.4 Acknowledgements 124
6.5 Note 124
7 GENERAL DISCUSSION 125
7.1 SIG5-dependent signalling pathway 125
7.2 But the world is complicated 126
7.3 Seasonality 128
7.4 Evolution of the temperature entrainment to the circadian clock 129
7.5 GENERAL CONCLUSIONS 130
8 APPENDIX 131
8.1 APPENDIX A: Correlation plots 131
8.2 APPENDIX B: Model script 133
8.3 APPENDIX C: Primers 134
8.4 APPENDIX D: Publications 136
9 REFERENCES 141

TABLE OF FIGURES

Figure 1 1. Diagram of the chloroplast structure……...7

Figure 3-1. Conserved σ70 family domains in the six sigma factors……...31

Figure 3-2. Multiple protein sequence alignment of AtSIG proteins …….....34

Figure 3-3. Electrostatic charges profile of the ancestral sigma factor from E.coli……35

Figure 3-4. Sigma factor RpoD in E. coli has positively-charged regions……...36

Figure 3-5. Three dimensional predicted model for SIG1……..37

Figure 3-6. Three dimensional predicted model for SIG2……..38

Figure 3-7. Three dimensional predicted model for SIG3……..39

Figure 3-8. Three dimensional predicted model for SIG4……..40

Figure 3-9. Three dimensional predicted model for SIG5……..41

Figure 3-10. Three dimensional predicted model for SIG6……42

Figure 3-11. Class I transcription activation complex……43

Figure 3-12. DNA-sigma factors models for Arabidopsis sigma factors……..44

Figure 3-13. Experimental and predicted phosphorylation sites in SIG1 and SIG2…..45

Figure 3-14. Predicted phosphorylation sites in SIG3 and SIG4. ……...46

Figure 3-15. Predicted phosphorylation site in SIG5 and SIG6. ……..47

Figure 4-1. Transcriptome data identifies SIG5 as highly responsive to cold……54

Figure 4-2. SIG5 is induced strongly in response to LT……54

Figure 4-3. The circadian oscillator gates LT signalling by SIG5……56

Figure 4-4. SIG5 is required for low temperature response in chloroplasts ……..57

Figure 4-5. HY5 and HYH act redundantly in the cold induction of SIG5……59

Figure 4-6. Low temperature signalling by SIG5 is HY5/HYH-dependent……..60

Figure 4-7. Abundance of four transcripts associated with cold adaptation……...62

Figure 4-8. SIG5 regulates the accumulation of SFR2……..63

Figure 4-9. Nuclear-encoded chilling responsive genes are regulated by SIG5……65

Figure 4-10. Nuclear-encoded chilling responsive genes are regulated by SIG5…….66

Figure 4-11. SIG5 positively regulates photosynthetic efficiency at LTs……..68

Figure 4-12. Electrolyte leakage……..……69

Figure 4-13. SIG5 is required for freezing tolerance…….……..70

Figure 4-14. Alterations in total triacylglycerol content (TAG) ……72

Figure 4-15. Fatty acid methyl ester (FAME) composition……...……..73

Figure 4-16. Accumulation of MEcPP but not SA induces SFR2……76

Figure 4-17. Schematic representation of the model proposed……..……77

Figure 5-1. Novel luciferase transient expression assay……..……81

Figure 5-2. SIG5 promoter from A. thaliana and A. halleri have the same activity…...82

Figure 5-3. Patterns of light and temperature recorded at the natural habitat……..….84

Figure 5-4. SIG5- mediated signalling to chloroplasts in the natural habitat…….……..85

Figure 5-5. Local environment manipulation equipment and patterns……86

Figure 5-6. Transcript abundance within the SIG5 pathway ……87

Figure 5-7. CCA1 peak time depends on light conditions……88

Figure 5-8. SIG5 transcript accumulation pattern……..……89

Figure 5-9. psbD BLRP transcript accumulation pattern……..……90

Figure 5-10. Plot of the natural logarithm of the components……..……..91

Figure 5-11. Evaluation of different delays between CCA1 and SIG5……94

Figure 5-12. Evaluation of different delays between SIG5 and psbD BLRP……..95

Figure 5-13. Environmental input terms improved model quality……98

Figure 5-14. Model failed to predict natural dataset ……..….99

Figure 5-15. Evaluation of models where SIG5 is uncoupled from CCA1…….……...100

Figure 5-16. SIG5 transcript abundance explained by CCA1……101

Figure 5-17. Major environmental drivers of the signalling pathway……102

Figure 5-18. Comparison of SIG5 experimental and predicted transcripts……..…103

Figure 5-19. In the natural dataset no environmental term improved model……..….104

Figure 5-20. Comparison of SIG5 experimental and predicted transcripts…….…….104

Figure 5-21. Inclusion of temperature improves the prediction of psbD BLRP…...….105

Figure 5-22. Inclusion of temperature improves the prediction of psbD BLRP…….....105

Figure 5-23. Inclusion of temperature improves the prediction of psbD BLRP…….....106

Figure 5-24. Environmental contribution to the model by R:FR……....107

Figure 5-25. psbD BLRP prediction is improved by FR…….…….….108

Figure 5-26. Blue light links the three components of the signalling pathway…….109

Figure 6-1. Morphology of Aegagropila linnaei…….114

Figure 6-2. Position within a body of water in the light period……..…..….115

Figure 6-3. Photosynthesis drives buoyancy of marimo…….…….116

Figure 6-4. Buoyancy under a blue/red light mixture is circadian-gated……..…….117

Figure 6-5. Chlorophyll fluorescence parameters Y(II) and ETR…….....118

Figure 6-6. Marimo rapid light response curves……..….119

Figure 6-7. Delayed fluorescence…….....….120

Figure 7-1. Modern closest relatives of land plants…….…….129

Table 2-1. Uniform gene nomenclature identifiers used in this study……..17

Table 4-1. Pathway enriched components in the co expression gene list…….…….75

Table 5-1. Time delays (h) between the two components of the models……..….93

Table 5-2. Correlation table between the eight environmental parameters …….…….97

Table 6-1. Y(II) and Y(NPQ) mean values……..…..….120

LIST OF ABBREVIATIONS

The following table describes the significance of various abbreviations and acronyms used throughout the thesis.

Abbreviation	Meaning
σ	Sigma
AIC	Akaike information criterion
CBF	C-REPEAT BINDING FACTOR
CCA1	CIRCADIAN CLOCK ASSOCIATED1
DCMU	3-(3,4-dichlorophenyl)-1,1-dimethylurea
DL	Dark/light
ETR	Electron transport rate
FR	Far red light
FFT NLLS	Fast Fourier transform-nonlinear least-square
Fv/Fm	Maximum quantum efficiency of Photosystem II
K_BT/e	Thermal energy (K_BT) normalized by charge (e)
LD	Light/dark
LL	Continuous light
LT	Low temperature
LUC	Luciferase
MESA	Maximum Entropy Spectral Analysis
ND	Neutral density filter
PAR	Photosynthetically active radiation
pKa	Acid dissociation constant
PSI	Photosystem I
PSII	Photosystem II
psbD BLRP	*psbD* blue light-responsive promoter
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
R	Red light
RAE	Relative amplitude error
R:FR	Red to far-red ratio
ROS	Reactive oxygen species
SA	Salicylic acid
SIG5	SIGMA FACTOR5
SFR2	SENSITIVE TO FREEZING2
TAG	Triacylglycerol
UV-B	Ultraviolet-B light
WT	Wild-type
Y(II)	PS II quantum yield during illumination
Y(NPQ)	Quantum yield of regulated energy dissipation
ZT	Zeitgeber time

1 INTRODUCTION

1.1 THE ARABIDOPSIS CIRCADIAN CLOCK

Due to the Earth's rotation on its axis, most life on Earth exists in day-night cycles of approximately 24 hours. Circadian clocks are timing mechanisms that help organisms to organize and coordinate their activities, and predict and adapt to fluctuations in their environment (Panda et al., 2002b). A key characteristic of circadian rhythms is that they are self-sustaining, which means they can persist in the absence of environmental cues. Circadian rhythms in plants regulate many processes such as photosynthesis (Dodd et al., 2015;Dodd et al., 2005), growth (Hotta et al., 2007;Green et al., 2002), response to stress (Kreps et al., 2002), flowering time (Yanovsky and Kay, 2002) and pathogen-herbivore defences (Seo and Mas, 2015).

A simplified view of the circadian clock is that it comprises inputs, the central oscillator and circadian outputs (Bell-Pedersen *et al.* 2005). The input signalling pathways perceive and communicate information from the environment to the central oscillator and sets its phase. In plants, the central oscillator is a complex machinery comprising around 20 proteins that form transcriptional and translational feedback loops, which in turn regulate and control the phase and rhythm of gene expression through circadian output pathways (Johansson et al., 2015)(Johansson et al., 2015). Outputs of the clock include gene expression (Harmer et al., 2000), chromatin remodelling (Malapeira *et al.* 2012), post transcriptional regulation (Sanchez *et al.* 2010) and post translational modifications (Panda *et al.* 2002), that finally confer rhythms to physiological and biochemical processes such as stomatal opening (Chen *et al.* 2012), flowering (Suarez-Lopez *et al.* 2001), and photosynthesis (Dodd *et al.* 2015).

Recent evidence indicates that each plant cell has its own circadian clock, independently entrained by light (Muranaka and Oyama, 2016) but it has also been shown that they share information in a hierarchical fashion, from the top to the bottom and from vasculature to mesophyll cells (Endo et al., 2014;Takahashi et al., 2015;Millar et al., 1995).

One of the feedback loops in Arabidopsis is composed of two Myb-like transcription factors, CIRCADIAN CLOCK ASSOCIATED1 (CCA1) and LATE ELONGATED HYPOCOTYL (LHY), and a member of the family PSEUDO–RESPONSE REGULATOR

(PRR) known as TIMING OF CAB EXPRESSION1 (TOC1). CCA1 and LHY peak at dawn and repress the expression of the dusk-phased TOC1 by binding to an element in its promoter known as the evening element (*EE,* AAAATATCT), which is in turn activated by the EE-binding transcription factor REVEILLE 8 (RVE8) (Hsu et al., 2013;Farinas and Mas, 2011;Rawat et al., 2011). RVE8 was identified by purification of EE-binding proteins in plants and is part of a MYB-like transcription factors family consisting of 11 homologous proteins (Rawat et al., 2011). The other members of the RVE8 family, RVE4 and RVE6 act redundantly to maintain accurate timekeeping (Hsu et al., 2013), whereas RVE3 and RVE5 have minor roles in clock function (Gray et al., 2017).

TOC1 in turn represses the expression of *CCA1* and *LHY* via the transcription factor CCA1 HIKING EXPEDITION (CHE) (Pruneda-Paz *et al*. 2009) and expression of the evening complex (EC; EARLY FLOWERING 3 (ELF3), ELF4 and LUX ARRHYTHMO (LUX)) (Mizuno et al., 2014).

CCA1 and LHY also promote the expression of other members of the PRR family: *PRR7*, *PRR9* and *PRR5*. These genes have a sequential pattern of expression during the day and are part of the feedback by suppressing *CCA1* and *LHY* (Nakamichi *et al.* 2010).

Other proteins required for correct function of the clock include ZEITLUPE (ZTL) and GIGANTEA (GI), which are involved in blue light sensing. ZTL is a special case of a clock component because the transcript abundance remains constant during the day, but protein levels are rhythmic peaking at night (Kim et al., 2007). It was found that ZTL is part of the ubiquitin ligase complex involved in targeting TOC1 for degradation at night, and is stabilised by GI (Kim et al., 2003), which induces CCA1 and LHY and is required for sucrose sensing and input to the clock (Dalchau et al., 2011).

1.2 INTEGRATION OF CIRCADIAN AND ENVIRONMENTAL INFORMATION

1.2.1 Light

The endogenous plant circadian oscillator must be entrained every day to keep in phase with the environment. One of the strongest timecues (or *zeitgebers*) is light, which also fuels photosynthesis, so plants must have a way of monitoring it. This is achieved through families of photoreceptors: the blue/UV-A (ultraviolet A) light sensing cryptochromes and phototropins, the red to far-red light absorbing phytochromes and the UV-B (ultraviolet B) (UVR8) light photoreceptor (McWatters and Devlin, 2011).

CRYPTOCHROME2 (CRY2) senses low fluence rate light while CRY1 perceives higher light intensities. PHYTOCHROME A (PHYA) perceives low red-light intensity and PHYB, PHYD, and PHYE transduce higher fluences rates of light to the clock (McWatters and Devlin, 2011). Reverse genetic studies conducted in mutants of phytochromes and cryptochromes showed that these mutants have longer period circadian oscillations (Nagel and Kay, 2012)(Nagel and Kay, 2012).

Besides light intensity, light quality is also important for plants. The spectrum of light from the sun is formed from wavelengths from 250 to 900 nm. However, chlorophyll and carotenoids absorb light at a wavelength below 480 nm and between 550 and 700 nm. This means that plants reflect light between 480 and 550 nm (green in the visible spectrum) and light above the 700 nm (far red). If a plant is close to another plant, the reflected light will be perceived by the adjacent plant, initiating a process called shade avoidance. Shade avoidance causes developmental alterations that allow plants to overcome possible shading by surrounding vegetation and its main trigger is a low Red (R) to Far Red (FR) ratio (R:FR) (Franklin 2008).

By applying the same light pulse to Arabidopsis at different times of day, it was found that the phase of the rhythm shifted forwards or backwards depending on the time of day (also known as phase response analysis) (Covington et al., 2001). This indicates that the circadian clock resets and entrains at critical times during the day. Furthermore, this response depends on the wavelength of light input, as illumination with red light renders greater phase advances than with blue light (Ohara et al., 2015).

Photoreceptor expression itself oscillates during the day. For example, *CRY1* and *PHYB* expression peaks at midday whereas *PHYA* and *CRY2* peak in the afternoon, which suggest that sensing light is gated by the clock (Hotta *et al.* 2007). Additionally, a phytochrome-binding bHLH transcription factor named PHYTOCHROME INTERACTIVE FACTOR3 (PIF3) interacts with PHYB and induces *CCA1* and *LHY* transcription (McClung 2006). Moreover, *PIF4* and *PIF5* (both bind PHYB) transcript abundance is high at the end of the dark period and mutants in these genes exhibited abnormal shade avoidance gene expression in low R: FR, providing a possible mechanism for the gating of shade avoidance (Franklin 2008).

Regarding ultraviolet light, low intensity UV-B light entrains the circadian clock through UVR8 and COP1 (Fehér et al., 2011). UV-B also induces *ELIP1* (*EARLY LIGHT*

INDUCIBLE PROTEIN 1) and *PRR9* regulated by LUX, ELF3 and ELF4. Furthermore, sensitivity to UV-B light depends on time of day (Takeuchi et al., 2014).

Altogether, this information points to the conclusion that light through the circadian clock provides spatial and temporal information about the environment.

1.2.2 Temperature and circadian regulation

Two alternating temperatures that differ by just 4 °C are capable of entraining the circadian clock in plants (McClung *et al.* 2002). This demonstrates that temperature as an input is as important as light to synchronize the oscillator to a dynamic environment.

Mechanistically, the clock genes *CCA1, LHY, TOC1, PRR7* and *PRR9* can be set to their correct phase by thermocycles (Salome and McClung 2005). *PRR7* and *PRR9* are especially important because they allow correct entrainment of the clock by temperature but are not required for entrainment by light (Salome and McClung 2005). Recently, a detailed study established that temperature signals are integrated into the clock via the evening complex (*ELF3, ELF4* and *LUX*) and that the direct targets of this complex are *PRR7, PRR9, GI* and *LUX* (Mizuno et al., 2014;Penfield, 2008;Gould et al., 2006).

Another process that links the circadian clock and temperature is termed temperature compensation. In this, the clock remains rhythmic with approximately the same period across a physiological range of temperatures (Penfield 2008). In Arabidopsis, temperature compensation occurs from 12 and 27 °C and arises from interaction between *CCA1*, *LHY*, *TOC1* and *GI* to maintain robust oscillations. For example, increasing the temperature causes that *TOC1* and *GI* buffer temperature changes while *CCA1* is the one that compensates during cooling (Gould *et al.* 2006).

A widely studied adaptive response to environmental stress is cold acclimation. In this, plants exposed to non-lethal low temperature exhibit multiple genetic and metabolic responses that lead to freezing tolerance (Fowler et al., 2005;Seo and Mas, 2015). The best characterized genetic system in responses to LT and subsequent cold acclimation is the *C-REPEAT BINDING FACTOR (CBF)* cold response pathway. Exposing Arabidopsis to LT induces transcription of *CBF1, CBF2* and *CBF3* (previously known as *DEHYDRATATION RESPONSIVE ELEMENT BINDING1b (DREB1b), 1c* and *1a*). CBFs belong to the AP2/ERF domain family proteins which recognize *C-REPEAT/DEHYDRATION RESPONSE ELEMENTS (CRT/DRE)* contained in the

promoters of *COLD REGULATED (COR)* genes such as *COR15A* and *COR78* (also known as *RD29A* and *LT178*). Overexpression of these genes confers similar metabolic changes as cold acclimation, but without the need for LT exposure (Fowler *et al.* 2005).

CBF-COR genes are induced by low temperature mostly during the day rather than the night (Fowler *et al.* 2005), having a clear circadian rhythm of expression considered to represent circadian gating. Gating is the term used when a response to environmental signals varies depending on the time of day the stimulus is given. This is thought to allow plants to perceive and respond to environmental signalling depending on the time of day, so that adaptive responses are most appropriate for the time of day (Seo and Mas 2015).

Increases in the concentration of cytosolic free calcium ($[Ca^{2+}]_{cyt}$) are one of of the first events that occur in response to LT in plants. Cold induced Ca^{2+} signals are gated by the circadian clock and correlate with increased expression of the cold regulated gene *COR15A* (Dodd et al., 2006). CCA1 and LHY bind to the *CBFs* promoters, increasing their expression, and *CCA1* and *GI* are induced by cold (Eriksson and Webb 2011). In contrast, TOC1, EC, and PRR9/7/5 proteins repress *CBF* transcription (Eriksson and Webb, 2011;Chow et al., 2014;Bieniawska et al., 2008). Recently, it was demonstrated that CBF1 binds the EE of the *LUX* promoter, regulating its transcription (Chow *et al.* 2014).

Other studies found that R:FR and temperature interact to coordinate responses to cold and acquisition of freezing tolerance. Plants treated with low R:FR at 16 and 22 °C exhibited gated *CBF* induction mediated by PHYB and PHYD (Franklin and Whitelam, 2007).

An extensive targeted expression analysis of circadian and cold induced genes under LT has provided information concerning the effect of cold in the clock (Bieniawska *et al.* 2008). Most circadian genes maintained their rhythmic expression but with lower amplitude than the controls when subjected to light/dark cycles. However, under continuous light, circadian oscillations are effectively stopped at 4 °C.

Temperature dependent transcript splicing has been shown to be important to integrate temperature and clock signals. The central oscillator gene *CCA1* can adopt two protein forms CCA1α that can bind DNA and CCA1β which is the inactive splicing variant. The abundance of each transcript form depends on temperature, thus low temperature leads to low *CCA1β* levels and higher active *CCA1α* contributing to freezing tolerance (Seo et

al., 2012). Other components of the circadian clock have been shown to be regulated by temperature dependent alternative splicing, such as *ELF3* and *TOC1* that in response to heat stress are targeted for nonsense-mediated decay due to intron retention (Kwon et al., 2014). Likewise, cold causes intron retention and subsequent degradation of alternative splicing forms of *ZTL* and *RVE8* (James et al., 2012). Recently, it was found that cold-responsive alternative splicing is widespread in the Arabidopsis transcriptome and occurs in a time-dependent manner (Calixto et al., 2018).

1.3 CHLOROPLAST GENOME

Pioneering work with lichens and amoeba in the beginning of the 20th century proposed that plastids were very similar to free living cyanobacteria (Mereschkowsky, 1920). However, the endosymbiotic theory was later proposed by Lynn Margulis after demonstrating DNA was present in the chloroplast (Plaut and Sagan, 1958).

Chloroplast genome sequencing suggest that these organelles conserve common genes with extinct photosynthetic bacteria and therefore emerged from a single endosymbiosis event between a cyanobacteria and a primitive eukaryotic cell through phagocytosis (Reyes-Prieto et al., 2007).

However, the protein importation machinery homologues Tic110 and Toc34 are found in plants and algae but are absent in cyanobacteria (McFadden and van Dooren, 2004) (Fig. 1-1), which contradicts the single endosymbiotic event theory. The current explanation is that once the plastid had established and endosymbiont genes migrated to the nucleus of the host, there was the need to acquire a system to import protein to the plastid followed by the evolution of the transit peptide (Reyes-Prieto et al., 2007).

Secondary endosymbiosis events occurred later, including the engulfment of a photosynthetic eukaryote from the red lineage that yielded cryptophyta, dinoflagellates and apicomplexans (Fast et al., 2001) whereas the common ancestor of euglenids and amoebas independently engulfed a cell from the green lineage (Rogers et al., 2007).

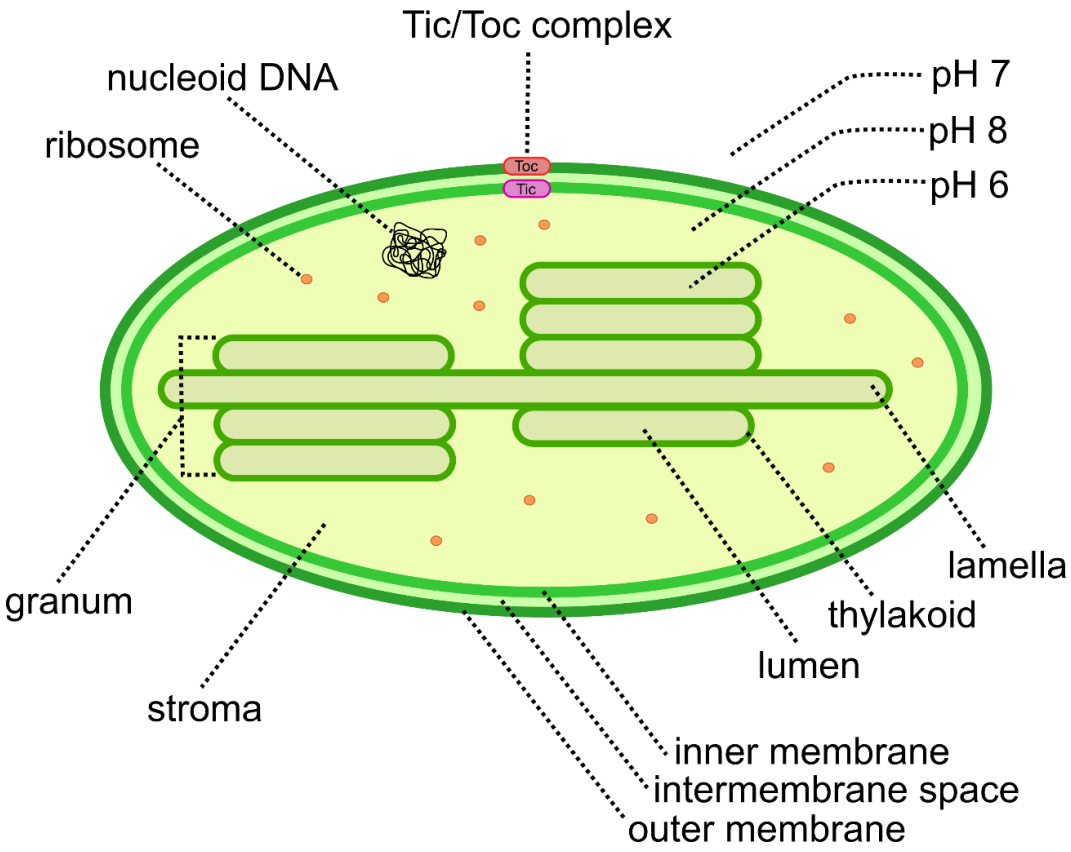


Figure 1-1. Diagram of the chloroplast structure. Chloroplasts are surrounded by a double membrane and contain thylakoids connected by the lamella. Stacked thylakoids are called grana or granal stacks (singular granum). The pH of the stroma and lumen are determined by the respective rates of electron transfer and ATP synthase activity in the thylakoid membrane. Import of nuclear-encoded proteins is through the Tic/Toc protein import complex.

In general, the chloroplast genome conserves part of the cyanobacterial-like genome, including components of the light reactions of photosynthesis and gene expression machinery. However, various regulatory components and most protein coding genes were transferred to the nuclear genome during evolution (Dorrell and Howe, 2012).

In organisms that have become parasitic and non-photosynthetic like the genus *Rafflesia* (Molina et al., 2014) or the protists Apicomplexa (Dorrell et al., 2014), the chloroplast genome has been greatly reduced or disappeared completely. This makes us wonder why photosynthetic organisms still retain many genes in their chloroplasts. One possible explanation is that those genes are transcriptionally regulated by redox feedback and therefore should remain within the chloroplast to allow rapid transcriptional regulation (Allen, 2015). Furthermore, the reason for transfer of chloroplast genes to the nucleus is still open, but plausible theories include avoiding damage from ROS produced by

photosynthesis, allowing sexual recombination or easier coordination with nuclear encoded genes (Dorrell and Howe, 2012).

1.4 CHLOROPLAST GENE EXPRESSION

The higher plant plastid genome is ~120-150 kb in size. In *Arabidopsis thaliana* it encodes 54 photosynthesis-related proteins, 31 proteins that regulate gene expression and 45 genes for tRNAs or rRNAs (Jarvis and Lopez-Juez 2013). It is important to consider that, as in prokaryotes, chloroplast genes are organized into operons. Most chloroplast encoded genes can be circadian regulated and there are also circadian rhythms of photosynthesis (Atkins and Dodd 2014). However, the mechanisms that drive the circadian oscillations of chloroplast gene expression is understood poorly.

Gene transcription in chloroplasts is executed by the nuclear encoded plastid RNA polymerase (NEP) and the plastid encoded plastid RNA polymerase (PEP). NEP is a T7 bacteriophage-type polymerase and is nuclear encoded. This enzyme transcribes housekeeping genes such as *accD*, and the *rpoB* operon. NEP was found to bind promoters containing the conserved YRT box consisting of pyrimidine (Y)-purine (R)-T designated Class I (Weihe and Börner, 1999).

PEP is homologous to bacterial RNA polymerase and consists of the core enzymatic subunits α, β, β', β'' (encoded by *rpoA*, *rpoB*, *rpoC1* and *rpoC2*, respectively) and one nuclear encoded sigma subunit responsible for promoter recognition (Nagashima *et al.* 2004). PEP is mainly involved in the transcription of photosynthesis related genes.

Higher plants contain various sigma factors that confer promoter specificity to the PEP machinery. In Arabidopsis, there are six sigma factors encoded in the nucleus (SIGMA FACTOR1 (SIG1)-SIG6). SIG2 and SIG6 participate in chloroplast development (Kanamaru 2001; Ishizaki *et al.* 2005), whereas SIG3 is responsible for the transcription of *psbN* (Zghidi et al. 2007) and SIG4 of the plastid *ndhF* gene encoding a subunit of the plastid NDH complex (Favory *et al.* 2005). Using chromatin immunoprecipitation techniques, the targets of SIG1 have been recently identified (Hanaoka *et al.* 2012).

SIGMA FACTOR5 (SIG5) induces the transcription of the *psbD* blue light-responsive promoter (*psbD* BLRP) in response to blue light and various environmental stresses including high light, high salt, low temperature and high osmotic conditions (Nagashima *et al.* 2004). *psbD* BLRP drives the transcription of *psbD*, which encodes the PSII

reaction centre protein D2. Unlike other promoters that have typical -10 and -35 elements that are recognized by sigma factors, *psbD* BLRP requires the -10 but not the -35 region. However, SIG5 can recognize this promoter by its -10 element and another conserved upstream enhancing sequence known as AAG box (Tsunoyama *et al.* 2004). SIG5 can also initiate transcription of *psbA*, which contains a functional -35 element (Onda *et al.* 2008). Also, it was found that SIG5 communicates timing information from the central oscillator to the chloroplast where it causes rhythmic transcription of *psbD* (Noordally *et al*. 2013).

Unlike bacterial genes, some chloroplast genes contain introns (Koller et al., 1984;Xu et al., 1990) which in the case of self-splicing processing in land plants requires nuclear encoded proteins (De Longevialle et al., 2010). Other genes are processed by splicing together two distinct RNA transcripts in a process denominated trans splicing which requires pentatricopeptide repeat (PPR) proteins to recognise and bind a sequence of 20 nucleotides around the coding region (Schmitz-Linneweber et al., 2006).

1.5 RETROGRADE SIGNALLING

As result of their endosymbiotic lineage most of the genes derived from the ancestors of chloroplast and mitochondria independently migrated to the nucleus of the eukaryotic cells. However, few of them remained inside plastids, making gene expression, photosynthesis and respiration machineries a mix of nuclear and plastic encoded proteins.

Communication between cellular compartments has emerged as an exciting new research field. Anterograde signals refer to nuclear encoded proteins that have a direct regulatory effect on plastid gene expression, whereas retrograde signals refer to signalling molecules coming from plastids that communicate the needs of the compartment to the nucleus (Kleine and Leister, 2016).

Early research in plants using blockers of plastid development such as norflurazon, which inhibits carotenoid biosynthesis (Oelmüller et al., 1986), or plastid translation inhibitors, lincomycin and chloramphenicol (Bradbeer et al., 1979), revealed that there were sets of photosynthesis-associated nuclear genes (PhANGs) controlled by the chloroplast. These are nuclear genes encoding components of the photosynthesis machinery (for example *LHCB1.2*) being transcribed even in the absence of a functional chloroplast (Oelmüller and Mohr, 1986). Using this approach, plants expressing GUS

under the control of *LHCB1.2* were developed and treated with ethyl methane sulfonate (EMS). There were five genomes uncoupled (*gun*) mutants identified in this original screen, with four found to be implicated in tetrapyrrole biosynthesis (*gun2* to *gun5*) (Susek et al., 1993). The main feature of the *gun* mutants is that they express PhANGs in the presence of the inhibitors, whereas a WT plant supresses these genes (Barajas-López et al., 2013).

Broadly, in WT plants tetrapyrroles synthesised from the precursor Glutamyl-tRNA can be chelated with $Fe^{2+/3+}$ or Mg^{2+} to produce porphyrins (haemes) or Mg-protoporphyrin IX (Mg-proto IX), respectively. Haem can be further modified to yield phytochromobilin, which is the light sensitive part of phytochromes whereas Mg-Proto IX converts to chlorophylls (Brunkard and Burch-Smith, 2018).

The signalling molecule by which tetrapyrrole biosynthesis has an effect on the nuclear genome has been fiercely debated. A first approach suggested GUN5 signalling involved the Mg-chelatase H subunit that is encoded by GUN5 (Mochizuki et al., 2001), then it was Mg-proto IX (Strand et al., 2003) which was eventually refuted (Mochizuki et al., 2008). The consensus is that GUN4 and GUN5 act as part of the Mg-chelatase complex, while GUN2 and GUN3 prevent the conversion of haem into phytochromobilin (Woodson et al., 2011).

Further research using the same battery of mutants identified *gun6*-1D, which is a mutant of the plastid ferrochelatase 1 (FC1, haem synthase) and it was suggested heme was the retrograde signalling molecule (Woodson et al., 2011). Further research showed that FC1-derived heme was used in all the cell whereas FC2 produces heme used only in the chloroplast and therefore is not part of a retrograde signal (Espinas et al., 2016).

Unlike the other gun mutants, *gun1* is a special case because in the presence of lincomycin the mutant still transcribes PhANGs (McCormac and Terry, 2004). GUN1 encodes a pentatricopeptide repeat protein (PPR) that has its effects in the nucleus through an Apetala 2 (AP2)–type transcription factor (ABI4) (Koussevitzky et al., 2007) or GOLDEN2-LIKE1 (GLK1) transcription factors (Waters et al., 2009). The exact mechanism of GUN1 is still not known, however it was suggested that it has its effects on nuclear transcription by the action of a membrane-anchored transcription factor in the chloroplast (TPM) (Sun et al., 2011). However, results from that study were not reproducible and later refuted (Page et al., 2017). It has also been reported that unlike other PPR proteins GUN1 has no interaction with nucleic acids but with proteins in the

tetrapyrrole synthesis pathway (Tadini et al., 2016). Hence, GUN1 has been proposed as a signal integrator of perturbances in the redox state, chloroplast gene expression, and tetrapyrrole synthesis (Kleine and Leister, 2016).

Light stress in chloroplasts generates reactive oxygen species (ROS) such as superoxide (O^{-}_{2}), hydrogen peroxide (H_2O_2) in the PSI and singlet oxygen ($^1O_2$) at PSII. The *fluorescent* (*flu*) mutant (Meskauskiene et al., 2001) accumulates protochlorophyllide, which is a photosensitizing agent that promotes $^1O_2$ production when exposed to light (op den Camp et al., 2003). The burst of $^1O_2$ has a signature response in nuclear encoded transcripts that differs from other ROS (Havaux, 2013). Further research into the *flu* mutants identified signalling components of this pathway including EXECUTER1 (EX1) and EX2 in the chloroplast (Wagner et al., 2004) and the nuclear PLEIOTROPIC RESPONSE LOCUS 1 (PRL1) (Flores-Pérez et al., 2010). However, further details about the operation of the pathway are not known.

ROS such as $^1O_2$ have further damaging effects in the chloroplast, for example, β-cyclocitral and dihydroactinidiolide are products from the oxidation of β-carotenes and can trigger EXECUTER-independent oxygen-responsive genes in the nucleus that partially overlap with those of *flu* mutants (Ramel et al., 2012).

In addition to ROS, plastoquinone reduction in response to high light revealed *ASCORBATE PEROXIDASE2 (APX2)* as a high light stress marker gene (Karpinski et al., 1997). Screening of EMS mutants of this marker gene promoter fused to LUC, revealed 13 *alx* mutant lines, among which *alx8* had higher levels of APX2 and ABA levels (Rossel et al., 2005). The *alx8* mutation was mapped to SAL1 which dephosphorylates PAP (3’ phosphoadenosine 5’ phosphate) to AMP (Wilson et al., 2009). SAL1 was shown to dimerize under oxidative stress, rendering it less active leading to PAP accumulation which causes inhibition in the activity of the 5’-3’ exoribonucleases (XRN) responsible for degrading stress inducible transcripts (Gy et al., 2007). PAP accumulates in response to drought and high light stress (Estavillo et al., 2011) but not cold, NaCl or ABA (Chen et al., 2011). Remarkably, it was recently described that osmotic stress prolongs the circadian period mediated by the SAL1-PAP pathway (Litthauer et al., 2018).

Another retrograde signalling pathway was discovered using a similar genetic screen of mutant plants expressing a stress-related marker gene. The nuclear gene *hydroperoxide lyase (HPL,* also known as *CYP74B)* encodes a protein involved in the breakdown of

fatty acid hydroperoxide in the chloroplast oxylipin pathway (Chehab et al., 2006). The promoter of *HPL* was fused to firefly luciferase and the resulting transgenic plants were EMS mutagenized. The screening identified 20 putative mutants with higher bioluminescence in the absence of environmental stress and the strongest bioluminescent line was selected and denominated *ceh1. CEH1* encodes 1-hydroxy-2-methyl-2-(E)-butenyl4-diphosphate synthase, an enzyme that catalyses the conversion of methylerythritol cyclodiphosphate (MEcPP) to hydroxymethylbutenyl diphosphate (HMBPP) in the plastidial methylerythritol phosphate (MEP) pathway. Further analysis of this line revealed that it contained ~4 fold more salicylic acid (SA) along with higher transcript levels of the nuclear-encoded *ICS1* involved in SA biosynthetic pathway. The signalling molecule was shown to be MEcPP, an isoprenoid precursor, because exogenous treatment of WT plants with this molecule induced *HPL* transcription and accumulated in response to wounding and high light stress in a *gun1*-independent fashion (Xiao et al., 2012).

1.6 LOW TEMPERATURE RESPONSES AND ADAPTION TO FREEZING IN THE CHLOROPLAST

For over a century it has been known that chloroplasts under bright light relocate as an avoidance response to light (Senn, 1909). More recently, a cold avoidance response has been discovered in cells of the fern *Adiantum capillus-veneris* alpine (Kodama et al., 2008) and *Arabidopsis thaliana* in which chloroplasts relocate in response to chilling (4 °C) (Fujii and Kodama, 2018) mediated by phototropin in *Marchantia polymorpha* (Fujii et al., 2017).

Furthermore sugars, such as sucrose, glucose, mannitol, sorbitol and others accumulate in freezing tolerant plants, lowering the freezing point and working as osmoprotectant agents (Kaplan et al., 2007). In response to cold, the β-amylase isoform from the chloroplast is more active, breaking starch into maltose that protects the PSII from cold damage measured as Fv/Fm (Kaplan et al., 2007).

Lipid remodelling processes have been widely characterised, because they prevent rupture of the plasma membrane, which has been identified as the first site of freezing injury (Steponkus, 1984). Besides adaptation of the photosynthetic machinery to low temperature conditions, it is known that chloroplasts also remodel their membranes to avoid rupture, which is lethal to the cell. COR15a is a CBF-dependent protein that accumulates in the chloroplast stroma during cold acclimation and it has been shown to

be a cryoprotectant protein by altering the curvature of the inner membrane of the chloroplast avoiding membrane fusion after freeze-thaw (Artus et al., 1996). COR15a counterpart in the outer envelope membrane of the chloroplast is SFR2, which has been characterised as a galactosyltransferase that uses monogalactosyldiacylglycerol (MGDG) as substrate, favouring lamellar membrane structure which avoids freezing injury (Thorlby et al., 2004;Moellering et al., 2010).

1.7 MOLECULAR BIOLOGY IN PLANTS UNDER NATURAL ENVIRONMENTS

Most of our knowledge about molecular systems in plants comes from laboratory-controlled settings. However, molecular pathways have been shaped to function under fluctuating and usually overlapping environmental stimuli (Kudoh, 2015). For example, how plants perceive environmental cues and filter noise are processes very poorly understood. In recent years, a novel combined approach to gain comprehensive information about molecular systems in plants has been developed. The term *in natura* was coined to describe knowledge from field studies and laboratory data (*in vivo* and *in vitro*) (Kudoh and Nagano, 2013).

This approach has been successfully applied to reveal the molecular mechanisms behind seasonal patterns in flowering time under natural conditions. FLOWERING LOCUS C (FLC) is a temperature-responsive repressor of flowering that is important for floral induction after prolonged cold periods (vernalization) (Michaels and Amasino, 1999). FLC suppresses the flowering switches *FLOWERING LOCUS T* (*FT*) and *SUPPRESOR OF OVEREXPRESSION OF CONSTANTS 1* (*SOC1*) (Searle et al., 2006). It was found that repression by FLC was released by exposure to low temperature mediated by histone modification (Bastow et al., 2004), specifically the trimethylation of histone H3 at lysine 27 (H3K27me3) in a time dependent way (Angel et al., 2011). Further studies showed that there are multiple processes leading to what has been called “winter memory” because after return to warm conditions, H3K27me3 mark is spread across the FLC locus depending on the length of the cold period (Angel et al., 2011).

The FLC homologue in *Arabidopsis halleri* subsp. *gemmifera* (*AhgFLC*) was studied in a natural population in central Japan. The study compiled *AhgFLC* expression, meteorological, and flowering data for over two years and analysis conducted by modelling revealed plants filter short term noise in temperature. The molecular machinery controlling flowering was found to extract seasonal patterns of temperature,

successfully ignoring flood, snow, drought and herbivory and short-term temperature fluctuations (Aikawa et al., 2010).

Transcriptomic tools and modelling have been also applied successfully to study rice in the field. Data from 461 microarrays from rice cultivars collected from May to October was analysed through linear modelling and researchers were able to predict 97% of genes expression solely based on meteorological data. This study revealed that other gene network systems are robust under natural fluctuating conditions and respond to few environmental cues (Nagano et al., 2012).

In field-based studies, the amount of data collected is considerably higher than under laboratory studies because sampling occurs in different seasons and high-resolution data is required to reveal patterns. Due to this, modelling has been incorporated to handle the amount of data and predictive models have been successfully used along *in natura* studies. For example, a predictive model for flowering in *A. halleri* revealed that global warming would cause significant decreases in flowering phase (Satake et al., 2013).

Building on these findings, a recent study carried out in field sites located in Sweden and UK found that besides requiring long term temperature memory, vernalization needs a period of daily temperatures below 15 °C (Hepworth et al., 2018).

1.8 PHOTOSYNTHESIS AND CIRCADIAN CLOCKS IN THE GREEN LINEAGE

The circadian clock of animals, plants, fungi and bacteria consist of a transcription-translation feedback loops, however, genes are not always orthologous, suggesting convergent evolution history (Serrano-Bueno et al., 2017).

In all plants, the central components of the clock *CCA1*, *LHY* and *PRR1 (TOC1)* have homologues in the members of the green lineage suggesting the core clock was acquired very early in evolution. However, other components such as *GI*, *ZTL*, *ELF3* and *ELF4* have no known homologues in micro algae and the moss *Physcomitrella*, but they are present in *Marchantia* and *Selaginella* suggesting these genes are exclusive to land plants (Linde et al., 2017).

Circadian rhythms of photosynthesis were first reported in enucleated *Acetabularia crenulata* cells using an oxygen sensor and electron transport inhibitors (Sweeney et

al., 1967), which challenged the notion that the circadian clock depends on transcriptional-translational feedback loops, because the oscillation appears to persist in the absence of the nucleus. Other examples of non-transcriptional oscillators are the KaiA-KaiC phosphorylation rhythms in *Synechococcus elongatus* and circadian rhythms of oxidation–reduction of peroxiredoxin proteins ubiquitous to most living organisms (Edgar et al., 2012). An interesting case of circadian rhythms orchestrating metabolic processes was found in nitrogen fixing cyanobacteria such as *Synechococcus* spp., that performs photosynthesis during the day and nitrogen fixation at night because the nitrogen-fixing enzyme nitrogenase is inactive in the presence of oxygen (Mitsui et al., 1986). As is the case for other flagellates, *C. reinhardtii* can swim towards the light (phototaxis) which is known to occur during the day while chemotaxis (to ammonium) occurs at night (Byrne et al., 1992).

Most of our knowledge about the circadian clock in algae comes from unicellular green algae such as *Chlamydomonas reinhardtii*, *Ostreococcus tauri* and *Acetabularia* that have been extensively studied. In the case of *C. reinhardtii* it has even been sent to space to study its circadian rhythm which was found to persist in the absence of gravity and magnetism (Noordally and Millar, 2015).

Circadian rhythms in macro algae have been described for the green marine alga *Bryopsis maxima*, by measuring oxygen levels under dark or light conditions. The oxygen depletion in darkness came from mitochondrial respiration while upon illumination, oxygen came from photosynthesis. It was found that there was a rhythm in oxygen evolution which persisted under continuous light (Okada et al., 1978). Circadian rhythms of photosynthesis have also been detected in 20 species of brown algae (Schmid et al., 1994), the marine red macroalga *Grateloupia turuturu* (Goulard et al., 2004) and the filamentous brown algae *Ectocarpus siliculosus* (Schmid and Dring, 1996).

The very few reports on circadian rhythms in macro algae mostly come from marine rather than freshwater algae. The importance of freshwater algae comes from the evolution of land plants, with their closest relatives coming from a strictly freshwater lineage, the charophyte green algae (Delwiche and Cooper, 2015). Therefore, it is important to study circadian regulation in freshwater algae as a basis for studying the evolution of circadian rhythms.

One of the very few examples of freshwater algae circadian rhythms come from studying *Volvox carteri*, which is a freshwater alga and one of the simplest multicellular organisms.

Interestingly, this organism is flagellated, and it has been reported that it swims at the end of the day towards the dark and cool areas and then at dawn swims back to the sunlight. Furthermore, transcripts of circadian clock genes accumulated in response to different light qualities (Kianianmomeni, 2014).

2 MATERIALS AND METHODS

In this methods chapter, I have organised the experimental procedures in the same order as the data chapters.

2.1 CHAPTER 3: THREE-DIMENSIONAL STRUCTURAL MODELLING OF ARABIDOPSIS SIGMA FACTORS

2.1.1 Multiple sequence alignment

Amino acid sequences from the six *Arabidopsis thaliana* sigma factors were retrieved from the TAIR11 database available at the TAIR webserver (https://www.arabidopsis.org/tools/bulk/index.jsp, accessed 08/2018) using the identifiers in Table 2-1.

Table 2-1. Uniform gene nomenclature identifiers used in this study.

Name	Model gene identifier	Locus
SIG1	AT1G64860	Chr1:24098497-24100746
SIG2	AT1G08540	Chr1:2703461-2706696
SIG3	AT3G53920	Chr3:19961041-19963820
SIG4	AT5G13730	Chr5:4429132-4430744
SIG5	AT5G24120	Chr5:8157794-8159746
SIG6	AT2G36990	Chr2:15537502-15540016

The sequences were then analysed for conserved domains using Pfam 31.0 (Finn et al., 2016), available at https://pfam.xfam.org/. The algorithm searches for conserved sequence domains.

Multiple domain identification and graphical output was obtained using PhylomeDB v4 (Huerta-Cepas et al., 2014) available at http://phylomedb.org/ (O24629 tree)

The PDB file from E. coli K12 RNA polymerase (RNAP) sigma70 holoenzyme was retrieved from the PDB website (https://www.rcsb.org/structure/6C9Y, accessed 08/2018). Chain F corresponds to the sigma factor template (sigf) to which to compare the *Arabidopsis thaliana* sigma factors three-dimensional structures.

Sequence alignments were performed, visualized and manipulated using Jalview 2.10.4b1 (Waterhouse et al., 2009), using the desktop version available to download at http://www.jalview.org/Web_Installers/install.htm. MUSCLE v3.8.31 (Edgar, 2004) with default settings was used as alignment logarithm and ZAPPO colour scheme was used with intensity set by conservation.

2.1.2 Homology modelling

The three-dimensional models for Arabidopsis sigma factors were created using the I-TASSER prediction server (Yang et al., 2014) available at https://zhanglab.ccmb.med.umich.edu/I-TASSER/, run with default settings. Briefly, after submitting the amino acid sequence to the server the Local Meta-Threading-Server (LOMETS) generates protein structure predictions that are used to find fragments of proteins with similar secondary structure from the PDB library. The resulting fragments are then reassembled, and the unaligned regions are modelled *ab initio*. The server then refines the structure using the TM-align structural alignment program to find the structurally closest PDB protein, and the restrains from LOMETS.

The native protein structures are identified using the SPICKER algorithm that removes steric clashes and defines a topology with the lowest free energy. Finally, the improvement of hydrogen bonds is performed by the REMO algorithm which also builds the final atomic model. Only one model was further analysed per sigma factor selected based on the highest estimated accuracy (C-score), where a higher value signifies a model with a high confidence, and a TM-score >0.5 which indicates a model with correct topology.

Once the models were selected, visualization and further analysis was made using PyMOL 1.7.4.5. Educational license (Schrodinger, 2015). Briefly, all structures were set to “cartoon” and the conserved domains detected previously were differently coloured to the rest of the structure.

2.1.3 pKa prediction

The pKa values of ionizable groups in amino acids were calculated using PROPKA 3.0 (Olsson et al., 2011) available through the PDB2PQR 2.0.0 server (more details below). Variations to the theoretical pKa values were calculated by considering the desolvation effects and intra-protein interactions which are empirically related to the position and interactions with other ionizable residues. All sigma factors predicted models were

submitted to the server and calculations were made at pH 7. The summary of the predictions was then normalised to neutral pH 7 to improve visualization of net charge. The results were then plotted against the amino acid position in the protein sequence. From the domain predictions, the approximate position within the sequence was indicated as boxes in the graph.

2.1.4 Macromolecular electrostatics determination

The predicted models for all sigma factors were submitted to the PDB2PQR 2.0.0 server (Dolinsky et al., 2004) available at http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/ (accessed 08/2018) with default settings. The server converts the occupancy and B-factor columns within a PDB file to charge and radius for each atom.

Macromolecular electrostatic potentials were calculated by the Adaptive Poisson-Boltzmann Solver (APBS) plug-in locally installed in PyMOL. This code enables to visualise the calculations as an electrostatic potential molecular surface. Colours were defined in a scale of +5 K_BT/e (blue) and -5 K_BT/e (red) and "Solvent accessible surface" and "Color by potential on sol. acc. surf." were selected to show only the potential in the surface that could be accessible to bind DNA.

2.1.5 DNA binding sites

Since the DNA binding domains are highly conserved between the six sigma factors and with the sigma70 from *E coli*, the experimentally determined DNA atomic model from the sigma factor in E. coli K12 RNAP holoenzyme (PDB ID 6C9Y) and the predicted models were aligned in PyMOL using the "align" tool and molecular clashes modified by "sculpting" tool.

2.1.6 Phosphorylation site prediction

Experimentally-identified and predicted phosphorylation sites were obtained from PhosPhAt 4.0 (Durek et al., 2010) available at http://phosphat.uni-hohenheim.de/phosphat.html (accessed 08/2018). Sigma factor identifiers submitted to the platform were the specified in Table 2-1.

The experimental phosphorylation site database was built from mass spectrometry data submitted by many research groups. The predictor was trained on the experimental database and sites are shown as coloured residues (S, T or Y). A phosphorylation

hotspot was defined as containing 4 phosphorylatable amino acids no further than 10 amino acids apart.

In PyMOL, experimental residues were then identified in the three-dimensional predicted model, selected and shown as yellow spheres. Predicted sites in hotspots were identified in the structures and shown as blue spheres. The rest of the predicted sites were shown as red sticks. Black arrows and labels were added in the figures for clarity.

2.2 CHAPTER 4: SIG5 COMMUNICATES INFORMATION TO THE CHLOROPLAST GENOME CONCERNING LOW TEMPERATURE CONDITIONS

2.2.1 Plant material and growth conditions

Seeds were surface-sterilized by treatment with an ethanol solution (70%, v/v) for one minute, followed by 10 min with a sodium hypochlorite solution (20%). Seeds were then washed three times with sterile water and resuspended in 0.1 % (w/v) agar and sown on half strength Murashige and Skoog basal salts mixture (Duchefa Biochemie) dissolved in 0.8% (w/v) agar. Stratification was performed in darkness at 4 °C for 2 days before transfer to Sanyo MLR-352 plant growth chambers. Cultivation occurred under cycles of 12 h light / 12 h darkness at 19 °C, 90 µmol m^{-2} s^{-1} of white light, with experiments starting at a seedling age of 11 days. Experiments used the T-DNA insertion mutants *sig5*-3, *sig5*-2 (Noordally et al., 2013; Nagashima et al. 2004) and *sfr2-3* (SLAK_106253) in the Col-0 background, and *hy5KS50 (hy5), hyh, hy5KS50 hyh (hy5 hyh)* and *cop1*-4 (Holm *et al.*, 2001) in the Wassilewskija (Ws) background. For inhibitor experiments, norflurazon (Sigma-Aldrich) was used at a concentration of 5 µM (0.1% (v/v) DMSO carrier) and lincomycin hydrochloride (Alfa Aesar) at 0.5 mM (dissolved in water) were added to growth media supplemented with 2% (w/v) sucrose to compensate the lack of nutrients result of chloroplasts rendered non-functional (Mochizuki et al., 2001).

For freezing tolerance experiments, seeds were stratified and imbibed for two days and then sown individually into plugs containing a 3:1 ratio of compost (Levington F2: horticultural silver sand (Melcourt)). Plants were grown under 110 µmol m^{-2} s^{-1} 12 h photoperiod light/dark cycles for 21 days at 19 °C and watered weekly. At week four, plants were exposed to -2 °C for 3 hours, commencing 1 h after dawn, then to -7 °C for 24 h, then -9 °C for one hour, before returning the plants to 4 °C for 24 h and then 19 °C (Hemsley et al., 2014). Plants were subsequently photographed, and survival was scored

21 days after the freezing treatment. Plants were scored as alive if the apical meristem was green or there was any green leaf tissue.

2.2.2 Electrolyte leakage assays

Quantitative determination of cellular damage after freezing was performed as described elsewhere (Hemsley et al., 2014) with some modifications. Three leaf disks per plant were excised using a cork borer and placed in cooled test tubes. Samples were rinsed by adding 5 mL of deionized water to each tube, shaking gently, then removing the water through a plastic mesh. Tubes were positioned upside-down on paper towels and left to dry for a few minutes. Afterwards, tubes were placed in a cooling bath to -2 °C for two hours and ice nucleation was induced by adding ice chips to the test tubes making sure the ice touched the leaves. After one hour, the cooling device was set to the first temperature to test (-3 or -7 °C for non-acclimated and acclimated plants respectively). Once temperature was stable, one-hour timing started. At the end of the one-hour treatment, three tubes per genotype per treatment were randomly removed from the bath and put on ice. The procedure was repeated for the next temperatures to test. Once all tubes were out of the bath, they were transferred to 4°C overnight and the following day 5 mL of deionized water was added to each tube. All tubes were then transferred to a shaker for 3 hours. At the end of the three hours, the liquid was transferred to a new clean set of test tubes and conductance was measured while the tubes with the leaves were put to -80 C for one hour to lyse all cells and release 100% electrolytes. Afterwards, the liquid was returned to the original tubes and shaked for another 3 hours. At the end, final conductance was determined, and percent of leaked electrolytes was calculated.

2.2.3 RNA extraction and qRT-PCR

Total RNA was isolated from frozen plant material using Macherey-Nagel Nucleospin II RNA extraction kits, and cDNA synthesized using an ABI High Capacity cDNA Reverse Transcription Kit as described previously (Applied Biosystems) (Noordally et al. 2013; Belbin et al. 2017). RNA concentrations were determined using a Nanodrop spectrophotometer (Thermo Scientific). 1/500 cDNA dilutions were analysed using Brilliant III Ultra-Fast SYBR Green QPCR master mix (Agilent Technologies) and appropriate primer sets (Appendix C), normalized to *ACTIN2* (Noordally 2013; Belbin 2017). For examination of transcript accumulation using conventional qRT-PCR, RNA isolation and cDNA synthesis was performed as above, and qRT-PCR conducted using DreamTaq (Thermo-Fisher).

2.2.4 Measurement of PSII photosynthetic efficiency

Chlorophyll fluorescence parameters were measured using an IMAGING-PAM MAXI chlorophyll fluorescence imaging system with pulse amplitude modulation (Walz GmbH, Effeltrich, Germany). Seedlings were cultivated for 11 days under light/dark cycles of 110 µmol m^{-2} s^{-1} white light, with a 12 h photoperiod, in 3:1 ratio of compost (Levington F2: horticultural silver sand (Melcourt)). Half of the plants were transferred to 4 °C for 10 days under light/dark cycles, at the same light intensity ("acclimated" treatment). The non-acclimated control comprised 21-day old plants grown at 19 °C. The third treatment, chilling, was provided by exposing 21-day old plants to 4 °C for 3h at under the same light conditions. For all experiments, plant pots were double-wrapped in aluminium foil under experimental temperatures to dark adapt for 60 minutes prior to fluorescence measurement. Measurements were initiated by exposing dark-adapted plants to measuring light pulses (1 Hz frequency, Intensity 3) and then applying a saturating pulse of 195 µmol m^{-2} s^{-1}. Fv/Fm was calculated as (Fm-F0)/Fm.

2.2.5 Lipid composition analysis

The lipid analysis was performed in collaboration with Tokiaki Takamura under the supervision of Dr. Sousuke Imamura at the Tokyo Institute of Technology, Japan. I prepared the samples, which were then shipped to Japan for chemical analysis.

Seedlings were cultivated as for qRT-PCR analysis. For cold acclimation, seedlings were placed at 4 °C for 10 days with the acclimation treatment starting at ZT 1. Non-acclimated controls remained at 19 °C. The freezing treatment comprised -2 °C for 3 hours, after which ice flakes were added to each plate to promote nucleation, and then seedlings were then cooled to -6 °C. Freezing treatments commenced at ZT 6 and lasted for at least 20 h. Harvested tissue was frozen in liquid N_2, freeze dried and stored in a protective atmosphere of Argon prior to analysis. For lipid extraction (Imamura et al. Plant Mol. Biol. 2015), about 50 mg of lyophilized tissue was combined with 600 µl of distilled water and 1 ml of chloroform:methanol (1:2, v/v) solution and homogenized using a mortar and pestle. The suspension was transferred to a test tube and mixed well with 2 ml of chloroform:methanol (1:2), and the supernatant was collected after centrifugation at 1000 g for 5 minutes at room temperature. The remaining material was mixed well with 3 ml of chloroform:methanol (1:2) solution and 800 µl of 1% (w/v) KCl solution, and a supernatant collected as above. The supernatant fractions were combined and mixed well with 3 ml of chloroform and 1.2 ml of 1% KCl solution. After centrifugation as above,

the lower phase containing lipids was transferred to a test tube and dried with nitrogen gas. After measurement of the weight of isolated total lipids, the total lipids were dissolved in 200 µl of chloroform:methanol (2:1, v/v) solution. TAG was isolated by a one-dimensional thin layer chromatography and analyzed by GC-FID, as described elsewhere (Imamura et al., 2015).

2.3 CHAPTER 5: SIG5-MEDIATED SIGNALLING TO CHLOROPLASTS UNDER NATURAL CONDITIONS

2.3.1 Plant material

Arabidopsis halleri subsp. *gemmifera* seeds were collected for laboratory work from one natural population located in Hyogo in central Honshu (along the Omoide River) in Japan.

Arabidopsis thaliana Columbia-0 (Col-0) expressing the *AhgpSIG5::LUC* construct were created as follows. The SIG5 promoter from *A. halleri* was amplified from genomic DNA by PCR using PstI-AhgpSIG5-F and BamHI-AhgpSIG5-R primers and digested with *PstI* and *BamHI*. The digested and purified product was cloned into a digested pGREENII0229::LUC plasmid (Noordally et al., 2013). The resulting plasmid (*AhgpSIG5::LUC*) was used for *Agrobacterium* mediated transformation in *A. thaliana* by floral dip and transformants were subjected to three cycles of selection by phosphinothricin resistance (BASTA) and bioluminescence assays.

2.3.2 Luciferase transient expression assay

Leaves from transformed plants resistant to BASTA (phosphinothricin) were vacuum infiltrated with 10 mL of 5 mM D-luciferin, potassium salt (Melford) and placed in a 24 well plate with MS agar. After 30 min, bioluminescence imaging was performed.

2.3.3 Bioluminescence imaging of promoter-luciferase reporter

AtpSIG5::LUC and *AhgpSIG5::LUC* plants of *Arabidopsis thaliana* were grown 11 days under 12 h LD cycles. 24 h before the experiment, plants were dosed with 100µL luciferin 5mM. Bioluminescence imaging was performed using a Photek HRPCS photon counting camera and data were analysed using Image32 software.

2.3.4 Measurement of light spectrum and temperature under field conditions

The light spectrum and intensity were measured using a USB2000+ spectrometer and a QP400-2-UV-VIS fibre optic cable attached to a CC-3-UV-S cosine corrector (Ocean Optics). Spectra were collected every 5 minutes over 14 hours of light during each season, for each light environment using the OceanView software (Ocean Optics) using a custom script. Power for the spectrometer and the computer was supplied by portable lithium battery packs (POWERGORILLA). Temperature was measured with a logger (EL-USB-2) every 5 minutes.

2.3.5 Field study time courses

A. halleri samples were collected from a natural population situated in Hyogo, central Honshu (along the Omoide River, Taka-cho; 35°06' N, 134°55' E, alt. 190–230 m) on two occasions: from the 24th to 26th of March 2015 and from the 15th to 17th of September 2015. In the field, two sites were selected to sample. In the site identified as sun, the plants received direct sunlight during the day while in the site identified as shade, plants received sunlight filtered by the vegetation most of the time. We sampled leaves from 6 plants randomly located at the sun site every 2 hours between 14:00 hours on the 24th and 13:30 hours on the 25th of March 2015 and between 14:00 hours on 15th and 13:30 hours on 16th of September. For the shade site, we collected leaves from 6 plants situated in different spots every two hours between 14:00 hours on 25th and 14:00 hours on 26th of March, and between 14:00 hours on 16th and 14:00 hours on 17th of September. At the same time the sampling occurred, the spectrum and temperature were measured near the plants.

In September 2016 we manipulated local conditions around patches of plants with the following treatments:

a) Control group, where the light meter was set to collect light spectra.
b) Warm conditions, which consisted of a clear plastic horticultural dome to block air currents and keep the heat.
c) Cool conditions. This included a device consisting of a fan connected to a pipe that went through a long ice box, ending in a clear horticultural dome above the group of plants. The system was aided by several small ice packs surrounding the plants. Every 8 hours, the ice box was refilled, and the packs changed.
d) Neutral density (ND) filter, low light conditions consisted on a sheet of ND filter wrapped around the plants and covering the top.

A control and under the ND filter spectra data were collected the1st September 2016 and %Transmittance was calculated taking the control light intensity as 100% at each wavelength. The ND filter spectra was then calculated from the control light spectrum in the field by multiplying the calculated %Transmittance.

To prevent missing points in the data and stress by excessive sampling damage, four spare plants were available at each site to sample. Also, to have a standard for qRT-PCR experiments, a pool of RNA from 10 leaves collected in March 2015 from random plants around the site was used to generate a reference cDNA to normalize all the samples from all sampling seasons.

All samples were shipped back to the UK and I performed RNA extraction, cDNA synthesis and qRT-PCR (*SIG5*, psbD *BLRP* and the clock gene *CCA1*) taking as reference the pool of samples collected in March 2015.

2.3.6 Preparation of RNA and qRT-PCR from *A. halleri* samples

After shipping of the samples in dry ice from Japan, samples containing RNA later were defrosted in the cold room for 4 hours and tissue was transferred to new dry tubes and refrozen in liquid nitrogen. Samples were kept at -80 °C until processing.

RNA extraction, cDNA synthesis and qRT-PCR were performed as described previously (section 2.2.3). RNA concentrations were determined using a Nanodrop spectrophotometer (Thermo Scientific) and RNA integrity assessed using a Bioanalyser (Agilent).

2.4 CHAPTER 6: TO FLOAT OR NOT TO FLOAT

2.4.1 Plant material and growth conditions

Because naturally-occuring marimo are extremely rare, all experiments were conducted using commercially-available marimo as an experimental model (k2aqua, purchased through Amazon UK). Marimo were washed thoroughly under cold tap water and stored in a large volume of water under 12 h light / 12 h dark cycles at 15 °C, under 10 - 15 µmol m^{-2} s^{-1} of white light. Immersion water from the tap was changed weekly and marimo were washed at regular intervals.

2.4.2 Microscopy

Images of dissected marimo balls were obtained using a Keyence VHX-1000 digital microscope at magnifications of between 5x and 200x. Dissection was performed using scissors and tweezers and mounted in petri dishes.

2.4.3 Buoyancy assay

Marimo were put in graduated measuring cylinders containing 500 ml tap water. Cylinders containing marimo were transferred to a plant growth chamber set to 15 °C and maintained under dim monochromatic red light (5 µmol m^{-2} s^{-1}) for 24 hours. Monochromatic blue light (15 µmol m^{-2} s^{-1} - providing a total of 20 µmol m^{-2} s^{-1} with the red light also) was switched on at either ZT 1 or ZT 7. Time-lapse images of the marimo-containing cylinders were acquired using an iPod (Apple, USA) with the Lapse It software application (v2.52, Interactive universe creative softwares EIRELI) at one minute intervals for 48 h, and time at which buoyancy was acquired was the difference between the start of blue light exposure and the time when the marimo buoyancy caused it to reach the 500 ml graduated line on the measuring cylinder. For buoyancy experiments using 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU, Sigma-Aldrich), the inhibitor was dissolved in dimethylsulfoxide (DMSO) to a 0.2 M concentration. DCMU solution was added to the water in the cylinders at the beginning of the dim red light period to a final concentration of 20 µM DCMU (DMSO, 0.01% v/v).

2.4.4 Measurement of PSII photosynthetic efficiency

Chlorophyll fluorescence parameters were measured using a Junior-PAM or IMAGING-PAM MAXI chlorophyll fluorescence system with pulse amplitude modulation (Walz GmbH, Effeltrich, Germany). For experiments under light/dark (LD) cycles, the Junior-PAM (WinControl-3 software) was coupled to an optical fibre that was positioned into the balls in the centre, halfway or outside the balls. Half of the balls were transferred to dark/light (DL) cycles for at least 5 days before measuring chlorophyll fluorescence. Experiments were initiated at ZT 3 for balls under LD cycle (day) and ZT 15 for balls under DL cycle (night) by wrapping individual balls in aluminium foil, placed in a labelled 5 cm petri dish to dark adapt them for at least 30 min at 15 °C. Measurements were initiated by exposing dark-adapted plants to measuring light pulses and then applying a saturating pulse (845 µmol m^{-2} s^{-1}), followed by blue actinic light (625 µmol m^{-2} s^{-1}) for two minutes.

To obtain light response curves, the actinic light intensity was increased in nine steps following each other within 10 s (0, 25, 45, 66, 90, 125, 190, 285, 420 µmol m^{-2} s^{-1}) and electron transport rate (ETR) was extracted from the software. Measurements were started from the interior of the ball, working outwards, with the optical fibre held in the same position for each depth of measurement within the marimo.

Circadian timecourses of rapid light response curves were obtained using the IMAGING-PAM MAXI (ImagingWin software) with marimo half-submerged in a tray. Marimo under light/dark cycles were cut in half at ZT 0 and rinsed under cold tap water. Half-marimo were placed in a tray with enough tap water to cover them, with either the surface or interior of the marimo facing the LED sources and camera equipment. Data were extracted from 4 regions of interest for each replicate marimo.

A custom instrument control program was used to acquire a timecourse of light response curves. Every two hours, plants were dark adapted for 20 min and measurements started by applying a saturating blue light pulse to obtain Fv/Fm. This was followed by a light response curve in which the actinic light was increased in 12 steps at 20 second intervals (0, 17, 33, 47, 73, 108, 150, 195, 245, 305, 369, 442 µmol m^{-2} s^{-1}). After acquiring the light response curve, the marimo were held under continuous blue light (15 µmol m^{-2} s^{-1}) until the next measurement cycle. For the circadian experiments, Y(II) and Y(NPQ) were determined from each light intensity of the light response curve, using the default Walz PAM configuration.

2.4.5 Delayed chlorophyll fluorescence analysis

Analysis of delayed chlorophyll fluorescence was based upon the approaches described elsewhere (Gould et al. 2009). Marimo balls were cut in half with scissors to expose the inside. Outside and inside halves were put in a black tray with enough water to cover them. Delayed chlorophyll fluorescence was measured using a Lumintek EM-CCD imaging system (Photek Ltd, St Leonards on Sea, UK) controlled by the Image32 software (Photek), under continuous red (24 µmol m^{-2} s^{-1}) and blue (29 µmol m^{-2} s^{-1}) lights. Images were captured at 30-minute intervals, over 96 h of constant light. Imaging used 45 second integrations (EM gain setting 3000). There was a 2 second pause between the lights switching off and onset of integration. Background-subtracted data were extracted from six regions of interests per marimo sample.

2.4.6 Quantitative timecourse analysis

For all timecourse analyses, the first 24 hours of data collected under continuous light were discarded and remaining data normalized to the first measurement in the timecourse for each region of interest. For delayed fluorescence analysis, a simple moving average every two data points was performed to reduce data noise (Gould et al. 2009). For PAM chlorophyll fluorescence analysis, Y(II) and Y(NPQ) values at each light intensity were obtained. Timecourses were analysed using the Biodare 2.0 platform (University of Edinburgh) (Zielinski et al., 2014), detrended by baseline and amplitude and analysed using the fast Fourier transform-nonlinear least-square method (FFT NLLS), with 18 to 30 h period cut-off. For each replicate marimo, the data were filtered to remove those regions of interest to which a cosine curve could not be fitted. A mean was obtained for each timepoint from the remaining regions of interest, and the resulting data analysed to obtain an estimate of period and quality of fit to the data (relative amplitude error, RAE). For PAM chlorophyll fluorescence analysis, the data were also analysed using the Maximum Entropy Spectral Analysis (MESA) algorithm (Biodare 2.0) as well as FFT-NLLS, to obtain a measure of rhythmicity using a different mathematical approach.

3 THREE-DIMENSIONAL STRUCTURAL MODELLING OF ARABIDOPSIS SIGMA FACTORS

The Pikachu effect

3.1 BACKGROUND

How DNA binding proteins, such as transcription factors, find their targets within millions of possible promoter sites to initiate transcription is a process that is poorly understood.

There are three proposed mechanisms for the task: sliding, intersegmental transfer and hopping (Yesudhas et al., 2017). Sliding refers to binding to non-specific regions in the DNA and sliding searching for specific sequence or promoter regions. Some proteins can diffuse to adjacent DNA strands without dissociating from the DNA which is known as intersegmental transfer. Hopping alludes to random collision events between the DNA and the protein until a stable interaction arises, however this process would be very energy expensive for the cell to repeatedly form and break multiple bonds between the protein and DNA (Halford and Marko, 2004).

Sliding and intersegmental transfer are regarded as the best strategies for DNA promoter recognition because they enable the protein to monitor multiple binding sites and diffuse to another strand without the need for dissociation from the DNA. Sequence specific interactions or “direct read-out” entail non-covalent interactions such as hydrogen bonds, electrostatic forces and van der Waals interactions (Rohs et al., 2010). Non-specific affinity for DNA or “indirect read-out” includes electrostatic interactions with the sugar-phosphate backbone of DNA and water molecules (Zhang et al., 2004). By using a combination of indirect and direct read-out activity, most DNA binding proteins find the specific sequence arrangements necessary for transcription initiation (Rohs et al., 2010).

Sigma (σ) factors are responsible for promoter and transcription initiation of RNA polymerase (RNAP) in archaea, bacteria, protists, some fungi, cyanobacteria, algae, moss and higher plants. Bacteria typically contain one σ factor for the RNAP activity in housekeeping genes (σ^{70}), however, they also can express alternative σ factors which are involved in spore formation (Johnson et al., 1983), flagellar gene expression (Arnosti and Chamberlin, 1989), synthesis of secreted products, ion uptake and stress responses (Helmann, 2002).

Members of the σ^{70} family have four conserved regions with specific activities. The σ^{70} region 1 is the most variable in composition and length. In bacteria this acts as an allosteric regulator of the enzyme by covering the DNA-binding regions when the factor is in solution but release the active site when attached to the core RNAP. The σ^{70} region 2 is the most highly conserved region and contains subdomains involved in binding to the core RNA polymerase, promoter melting and binding to -10 elements in promoters (TATAAT box). The σ^{70} region 3 is involved in the binding to the core RNA polymerase while σ^{70} region 4 recognises -35 DNA motifs (TTGACA consensus) (Paget and Helmann, 2003). The promoter motifs can be different depending on the function, for example σ^{70} involved in flagellar gene expression recognises GCCGATAT rather than the canonical TATAAT box in -10 elements recognised by σ^{70} region 2 (Arnosti and Chamberlin, 1989).

Gene transcription in chloroplasts is executed by nuclear encoded plastid RNA polymerase (NEP) and plastid encoded plastid RNA polymerase (PEP). NEP is a T7 bacteriophage-type polymerase and is nuclear encoded. This enzyme transcribes housekeeping genes such as *accD*, and the *rpoB* operon. PEP is a homolog to bacterial RNA polymerase and consists of the core enzymatic subunits α, β, β', β'' (encoded by *rpoA*, *rpoB*, *rpoC1* and *rpoC2*, respectively) and one nuclear encoded sigma subunit responsible for promoter recognition (Nagashima *et al.* 2004). This PEP is mainly involved in the transcription of photosynthesis related genes.

Higher plants contain various sigma factors that confer promoter specificity to the PEP machinery. In Arabidopsis, there are six sigma factors encoded in the nucleus (SIGMA FACTOR1 (SIG1)-SIG6). SIG1 is involved in PSI and PSII transcripts such as *psaAB*, *psbBT* and *psbEFLJ*, a Rubisco subunit (*rbcL*) and a subunit of the Clp protease (*clpP*) (Hanaoka et al., 2012). SIG2 and SIG6 participate in chloroplast development (Kanamaru 2001; Ishizaki *et al.* 2005) whereas SIG3 is responsible for the transcription of *psbN* (Zghidi et al. 2007) and SIG4 of the plastid *ndhF* gene encoding a subunit of the plastid NDH complex (Favory *et al.* 2005).

SIGMA FACTOR5 (SIG5) induces the transcription of the *psbD* blue light-responsive promoter (*psbD* BLRP) in response to blue light and various environmental stresses including high light, high salt, low temperature and high osmotic conditions (Nagashima *et al.* 2004). *psbD* BLRP drives the transcription of *psbD*, which encodes the PSII reaction centre protein D2. Unlike other promoters that have typical -10 and -35 elements that are recognized by sigma factors, *psbD* BLRP has the -10 but not the -35 region.

However, SIG5 can recognize this promoter by its -10 element and another conserved upstream enhancing sequence known as AAG box (Tsunoyama *et al.* 2004). SIG5 can also initiate transcription of *psbA*, which contains a functional -35 element (Noordally et al., 2013;Onda et al., 2008).

In this chapter I performed analysis to understand the biochemical differences between the sigma factor proteins from *Arabidopsis thaliana.* I performed this to gain insights into the basis for the specificity of transcription conferred by each sigma factor. This involved multiple sequence alignments, electrostatic force profiles, three-dimensional protein modelling and phosphorylation sites mapping.

3.2 RESULTS

First, I examined the amino acid sequences of all *A. thaliana* sigma factors and looked for conserved domains of DNA binding regions. Sequences were aligned with other members of the σ^{70} family using the Pfam webserver to identify conserved motif families (Finn et al., 2016) and the multiple domain identification software PhylomeDB (Huerta-Cepas et al., 2014). All 6 Arabidopsis sigma factor homologues were identified to have the conserved σ^{70} regions: 2,3 and 4 (Fig. 3-1).

Since the best and most widely characterised sigma factor is RpoD from *E. coli*, the protein sequence was retrieved from Uniprot and used for the following analyses to compare the Arabidopsis sigma factors against this subunit.

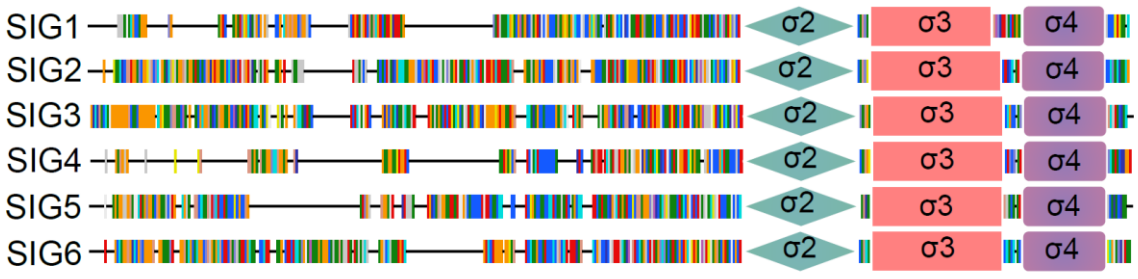


Figure 3-1. Conserved σ^{70} family domains were found in the six sigma factors from *Arabidopsis thaliana*. Output from the multiple domain identification from PhylomeDB. Motifs are represented by different shaped boxes. Inter-domain coding regions are shown using the standard amino acid colour codes. Gaps in alignment are illustrated as a flat line.

All sigma factors contain the same domains for recognition and binding of DNA, and since they were identified by sequence conservation alignments, I wondered whether

they were similar sequences or differed in amino acid properties. To investigate this, multiple sequence alignment was performed, and amino acids were coloured according to their physicochemical properties with intensity set by conservation (30% threshold). Conservation was determined based on the AMAS method of multiple sequence alignment (Livingstone and Barton, 1993). The highest sequence conservation is within the σ^{70} regions, especially promoter recognition domains 2 and 4 (Fig. 3-2).

Furthermore, in the conserved domains from sigma factors, there are many instances where a negatively charged amino acid in one sigma factor is positively charged or hydrophobic in other sigma factor (between positions 470 and 540 in Fig. 3-2) and vice versa. Due to this, I reasoned that higher plant sigma factors specificity might be influenced by the electrostatic properties of the amino acids. All amino acids contain an amino group, a carboxylic acid, a hydrogen and a sidechain. The charge of every amino acid depends on their pKa values and the pH they are exposed to. Proteins contain negatively (glutamate and aspartate) and positively (lysine and arginine) charged residues in addition to amino acids with local environment-dependent charge (histidine, cysteine and tyrosine). However, the real charge depends on the molecules around them and especially the buried coefficient, which depends on the position in the three-dimensional structure (Isom et al., 2011). Charges were calculated by estimating the pKa (acid dissociation constant) of the amino acids that constitute the sigma factors and were calculated at physiological pH 7 for all charged amino acids. To facilitate a visualization of the charge in the sequence, the values of pKa were normalised to neutral, so that a pKa of 8 is -1 and indicates a negative charge (electrons are available) whereas a pKa of 6 gives a value of +1 (positive net charge) which can then interact with the negatively charged DNA backbone. The resulting values were then plotted against the sequence as an area chart with the conserved binding regions 2, 3 and 4 indicated.In general, each sigma factor has a unique electrostatic pattern (Fig. 3-3). Compared to the σ^{70} from *E. coli*, all plant sigma factors in the regions outside the conserved domains lack a positively charged region at around the 100 amino acid position (Fig. 3-3). SIG1 is enriched in negatively charged amino acids compared to the *E. coli* sigma factor and the other sigma factors in Arabidopsis (Fig. 3-3B). SIG2 lacks the negative patch in the beginning of the sequence (position 0 to position 100) found in most of the other plant factors (Fig.3-3B). SIG3 and SIG4 have a very similar charge pattern before the 200 amino acid position and within the σ^{70} conserved domains, however SIG4 seems to be the short version of SIG3 (Fig. 3-3D, 3-3E). SIG5 and SIG6 have very similar charges in the regions corresponding to the σ^{70} regions 2 and 3 (Fig. 3-3F, 3-3G).

The results suggest there are differences in the electrostatic properties of each sigma factor. However, only the amino acids in the surface of the protein structures are in contact with DNA molecules, therefore, we mapped the charges onto the surface, and visualised the charges on the three-dimensional structures to investigate whether there might be electrostatic differences between the sigma factors that could contribute to their promoter specificity.

3.2.1 Three-dimensional homology modelling

All Arabidopsis sigma factors were submitted to I-TASSER (Yang et al., 2014) which generates protein structure models based on sequence homology to experimentally determined three-dimensional structures from the Protein Data Bank (PDB). The resulting five top models were then ranked by their quality scores and the best model was selected based on the highest estimated accuracy (C-score), where a higher value signifies a high confidence model, and a TM-score >0.5 which indicates a model with correct topology.

Once the models were selected, visualization and further analysis was made using PyMOL 1.7.4.5. Educational license (Schrodinger, 2015). Briefly, all structures were set to “cartoon” and the conserved domains detected previously (Fig. 3-1) were coloured differently from the rest of the structure.

The predicted models for all sigma factors and the experimentally determined sigma factor from *E. coli* were submitted to the PDB2PQR 2.0.0 server (Dolinsky et al., 2004) with default settings (using PARSE forcefield) to obtain the van der Waals radius and electrostatic charge data derived from the calculated pKas. Due to the slightly higher pH inside the chloroplast lumen where the sigma factors function, the pH was set to 8.

Figure 3-2. Multiple protein sequence alignment of AtSIG proteins and the homologue in *E. coli* using MUSCLE showing coloured amino acid residues (ZAPPO palette) and intensity set by conservation. A quantitative scale for conservation (0-9,+, *) is shown as the histogram row in yellow. + equals to 10 and indicates there was a mutation, but all properties are conserved. ***** indicates same amino acid identity. The conserved regions are shown as boxes.

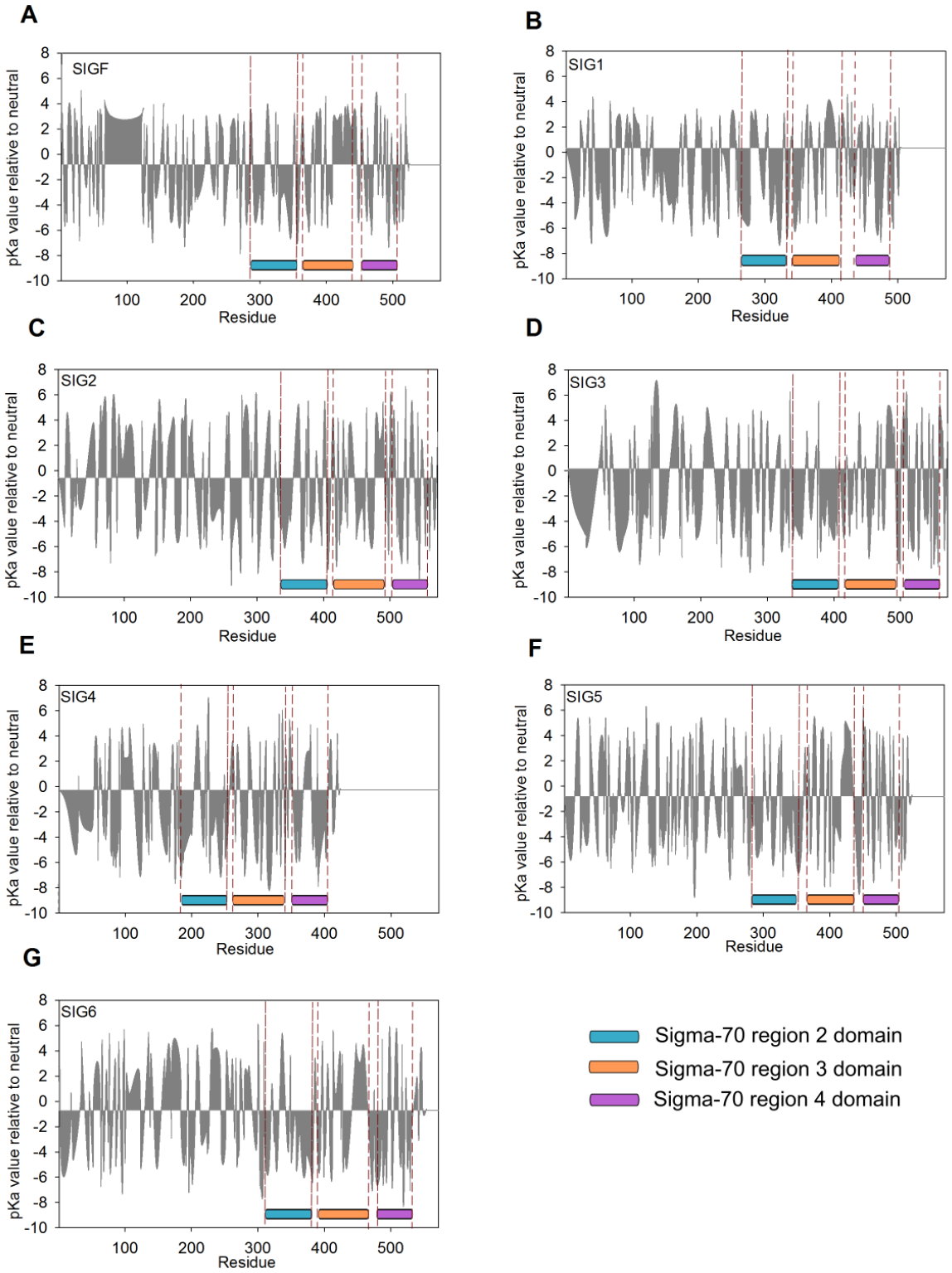


Figure 3-3. Electrostatic charges profile of the ancestral sigma factor from E.coli and amino acid sequences from the six Arabidopsis thaliana sigma factors. Calculated pKa are shown relative to pH 7 and the conserved regions are shown with boxes along dotted lines for clarity.

Macromolecular electrostatic potentials were calculated by the Adaptive Poisson-Boltzmann Solver (APBS) plug-in locally installed in PyMOL. This code enables visualization of the calculations as an electrostatic potential molecular surface. Colours were defined in a scale of +5 K_BT/e (blue) and -5 K_BT/e (red) and volume set by solvent accessible surface to show only the potential that could be accessible to bind DNA.

The contact surface of DNA binding proteins is characterised by high positive electrostatic potential which might help to steer the DNA leading to conformational changes in the protein and DNA (Marcovitz and Levy, 2011). A clear example of this is shown in the sigma factor from *E. coli* (PDB ID 6B6H) where the σ^{70} regions 2 and 4 have distinctive positively charged patches that correspond to -10 and -35 motif binding regions (Fig. 3-4). Interestingly, there is another highly positive patch in the surface of the protein (marked in dotted yellow lines, Fig. 3-4A).

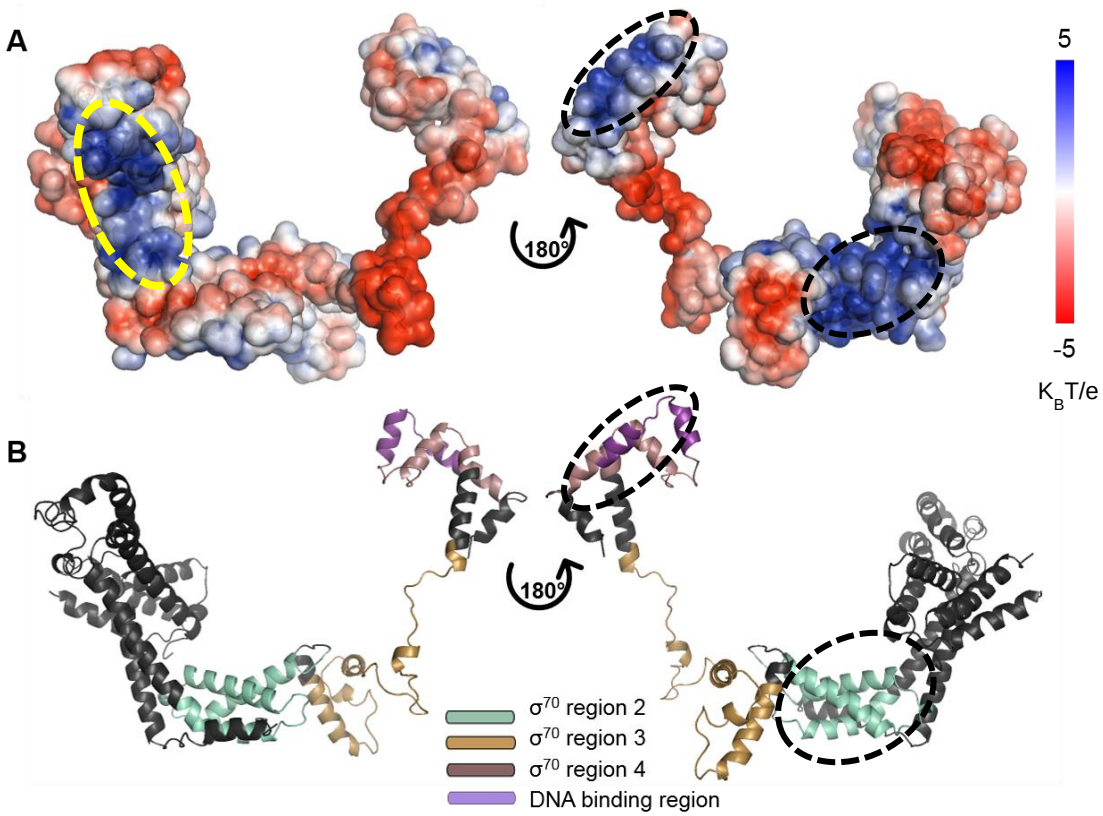


Figure 3-4. Sigma factor RpoD in *E. coli* has positively-charged regions. A) Surface charge distribution of the structure of a bacterial class I transcription activation complex (PDB ID 6B6H chain F). The positive patches (blue, circled) correspond to regions 2 and 4 in B) the helix-coil representation of the protein.

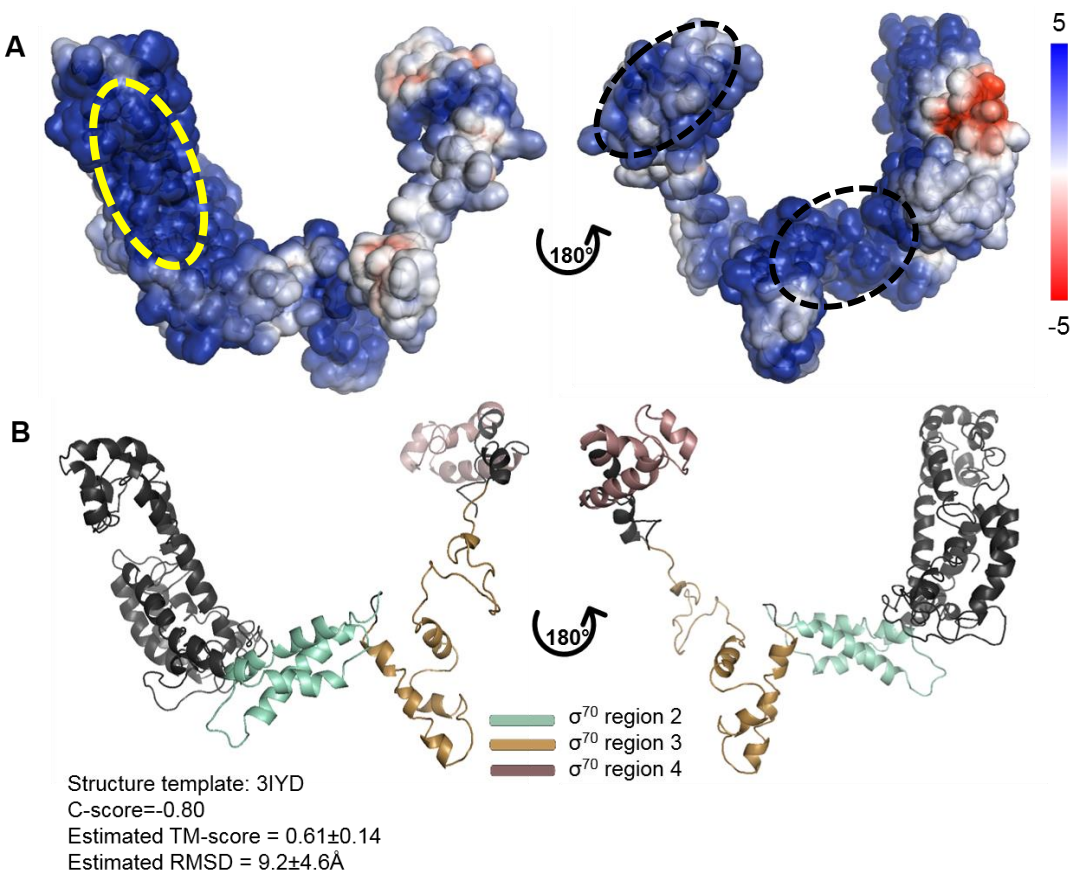


Figure 3-5. Three dimensional predicted model for SIG1 with surface charge distribution (A) and conserved σ^{70} regions (B). Electrostatic surface potential scale was set to +5 K_BT/e (blue) and -5 K_BT/e (red). Patches corresponding to positively-charged regions in RpoD from *E. coli* (Fig. 3-4) are circled in dotted lines.

A SIG1 homology model was based on the activator-dependent transcription initiation complex from *E. coli* K12 that includes the RNA polymerase holoenzyme (RNAP) (PDB ID 3IYD). Unexpectedly, most of the protein surface is positively charged, with very few negative regions except for a patch outside the conserved σ^{70} regions and neutral areas within the σ^{70} region 3 (Fig. 3-5).

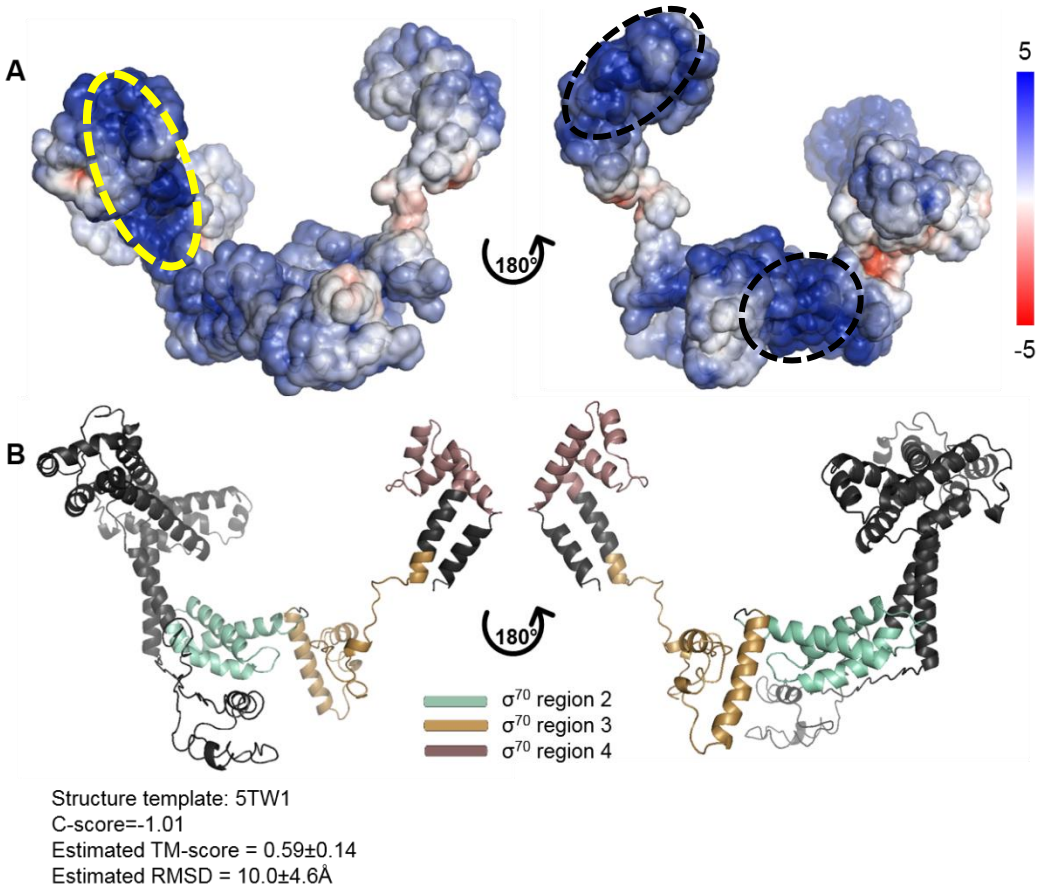


Figure 3-6. Three dimensional predicted model for SIG2 with surface charge distribution (A) and conserved σ^{70} regions (B). Electrostatic surface potential scale was set to +5 K_BT/e (blue) and -5 K_BT/e (red). Patches corresponding to positively-charged regions in RpoD from *E. coli* (Fig. 3-4) are circled in dotted lines.

A SIG2 three-dimensional structure was modelled based on the sigma factor from the transcription initiation complex from *Mycobacterium smegmatis* (PDB ID 5TW1). In general, SIG2 is relatively more neutral and less positive compared to the modelled SIG1 surface, but the DNA binding domains remain as positively-charged patches (circled with black dotted lines) (Fig. 3-6).

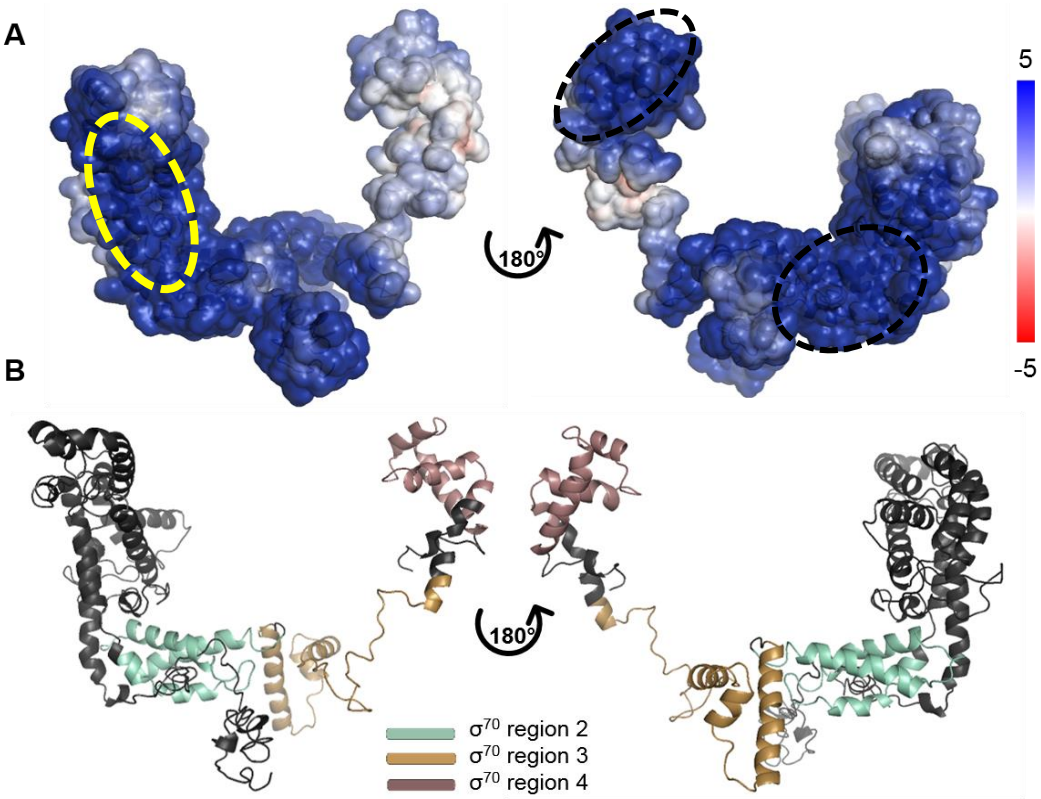


Figure 3-7. Three dimensional predicted model for SIG3 shown as surface charge distribution (A) and conserved σ^{70} regions (B). Electrostatic surface potential scale was set to +5 K_BT/e (blue) and -5 K_BT/e (red). Patches corresponding to positively-charged regions in RpoD from *E. coli* (Fig. 3-4) are circled in dotted lines.

SIG3 was also modelled based on the activator-dependent transcription initiation complex from *E. coli* structure, however, no negatively charged patch was found in this structure, only few neutral areas (Fig.3-7).

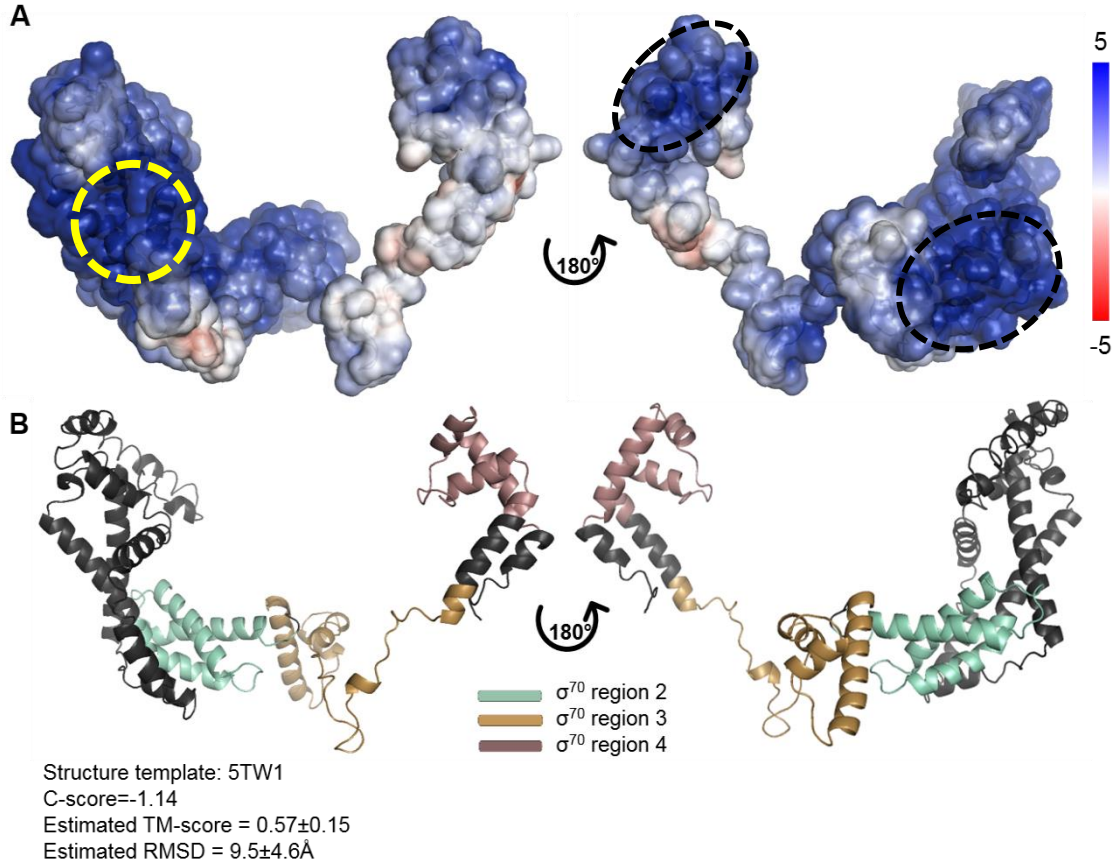


Figure 3-8. Three dimensional predicted model for SIG4 shown as surface charge distribution (A) and conserved σ^{70} regions (B). Electrostatic surface potential scale was set to +5 K_BT/e (blue) and -5 K_BT/e (red). Patches corresponding to positively-charged regions in RpoD from *E. coli* (Fig. 3-4) are circled in dotted lines.

SIG4 has the shortest amino acid sequence even when compared with the *E. coli* sigma factor, therefore it is not surprising the three-dimensional structure is also the smallest. In this modelled structure, the main difference compared to the other sigma factors is the short non-conserved region. However, the positively charged patch observed in all the other factors (circled in yellow dotted line) is present (Fig. 3-8). In general, the surface is mostly positively charged except for the areas that link σ^{70} regions 2 to 3 and 3 to 4 which are neutral.

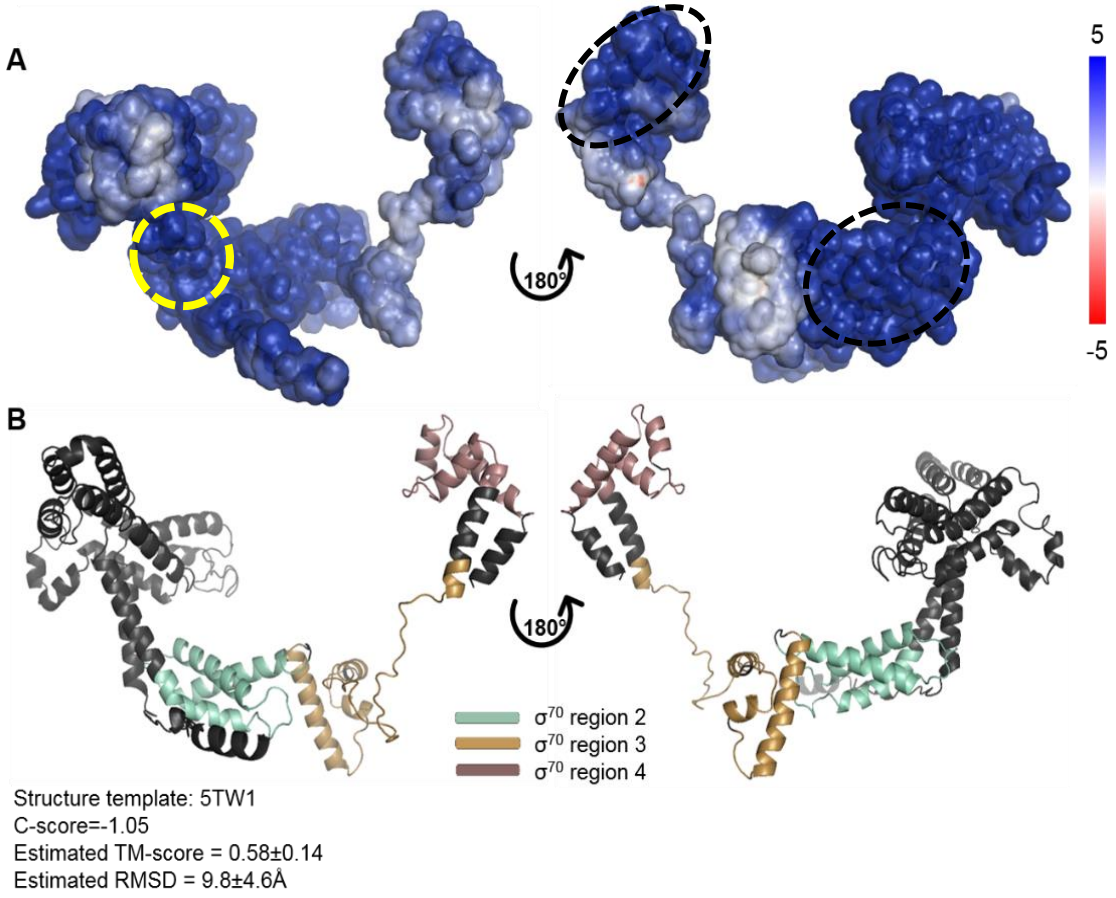


Figure 3-9. Three dimensional predicted model for SIG5 with surface charge distribution (A) and conserved σ70 regions (B). Electrostatic surface potential scale was set to +5 KBT/e (blue) and -5 KBT/e (red). Patches corresponding to positively-charged regions in RpoD from E. coli (Fig. 3-4) are circled in dotted lines.

SIG5 was modelled based on the *Mycobacterium smegmatis* sigma factor and contains a neutral σ^{70} region 3. Overall, the surface is very positively charged with a spot of negatively charged area (Fig. 3-9).

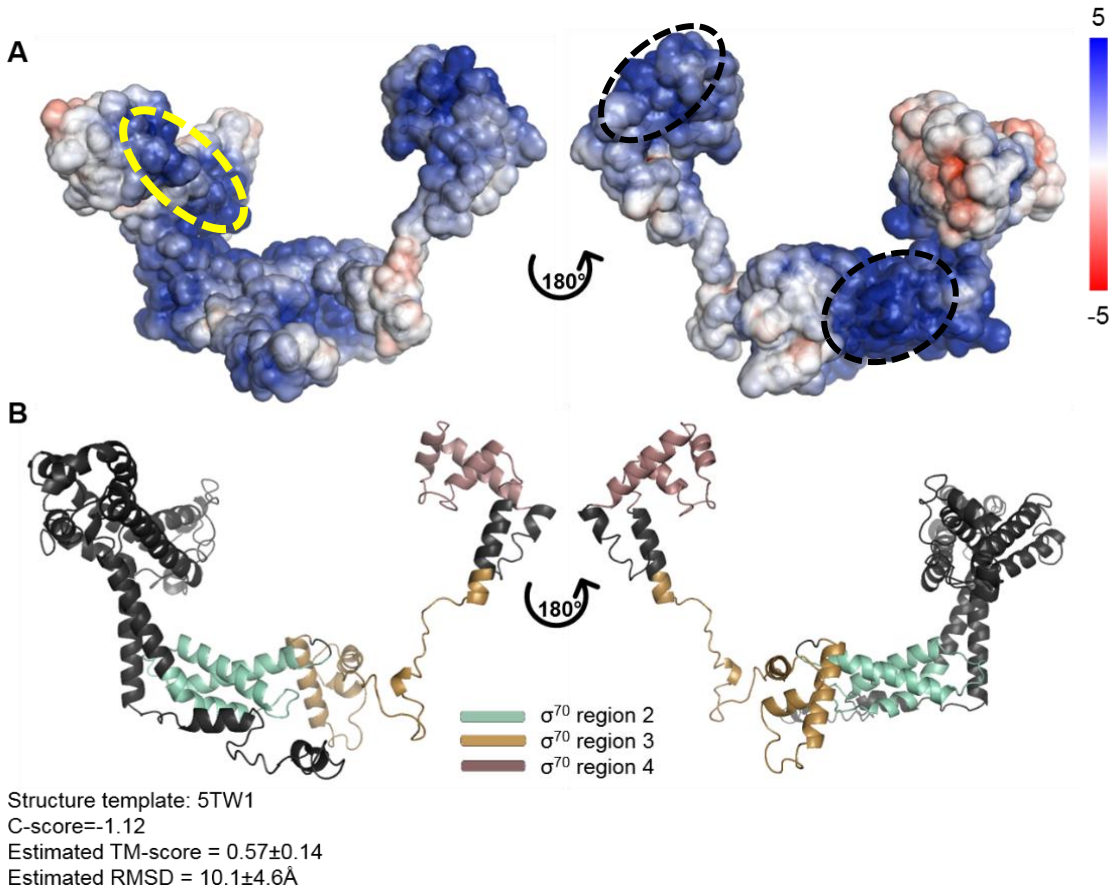


Figure 3-10. Three dimensional predicted model for SIG6 shown as surface charge distribution (A) and conserved σ70 regions (B). Electrostatic surface potential scale was set to +5 KBT/e (blue) and -5 KBT/e (red). Patches corresponding to positively-charged regions in RpoD from E. coli (Fig. 3-4) are circled in dotted lines.

SIG6 is the least positively charged of all plant sigma factors and it shows very clear DNA-binding domains that correspond to σ^{70} regions 2 and 4. Some of the negatively charged areas from the *E. coli* sigma factor are conserved, however the negative charge is weak (Fig. 3-10).

The positively charged patch near the N-terminal in the *E. coli* sigma factor (Fig. 3-4A, yellow circle) has no reported function due to its location outside the DNA-binding domain, however, it is also present in the plant sigma factors, so its function was interesting to us. Taking the complete holoenzyme RNAP structure from PDB and marking this zone in the sigma factor, I noticed that it was near a helix from another subunit. By identifying positively and negatively charged amino acids in this area, I identified a positive patch in the sigma factor that has an opposing charge region in the β' subunit from RNAP.

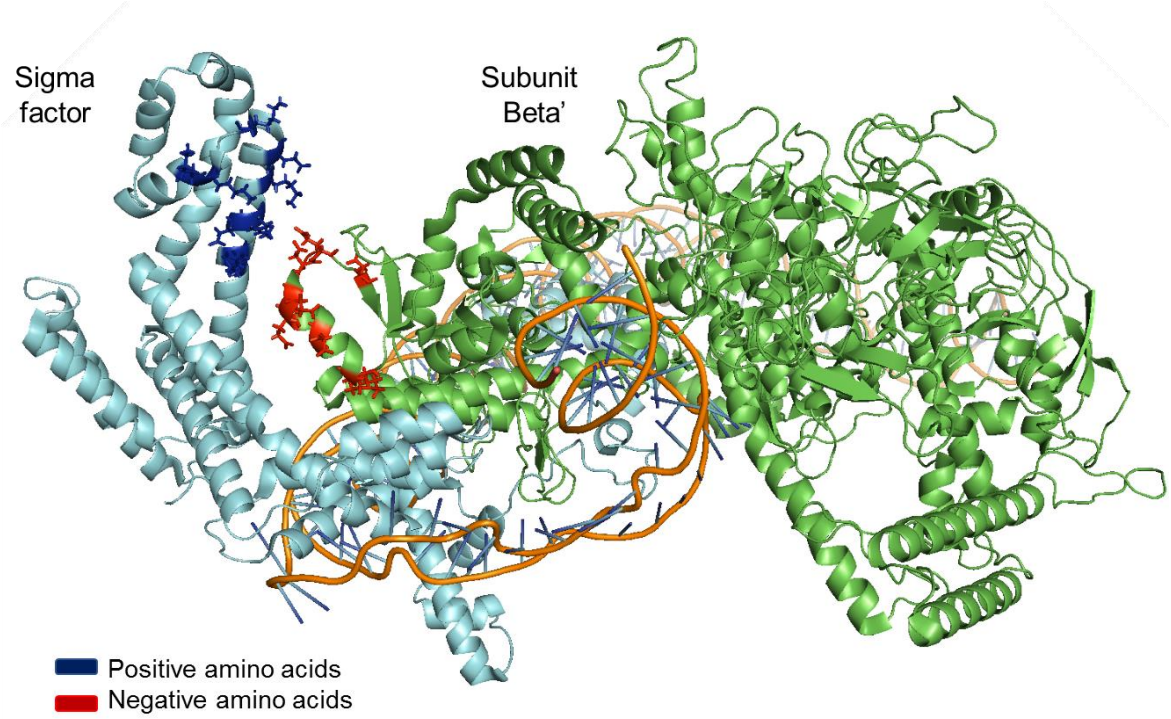


Figure 3-11. Extract from the bacterial class I transcription activation complex (PDB ID 6B6H) showing the suggested interaction between the positively charged area in the non-conserved segment from the sigma factor (sky-blue) and the negatively charged area found in the β' subunit (green). DNA is shown as an orange ribbon. Charged amino acids are shown as sticks for clarity.

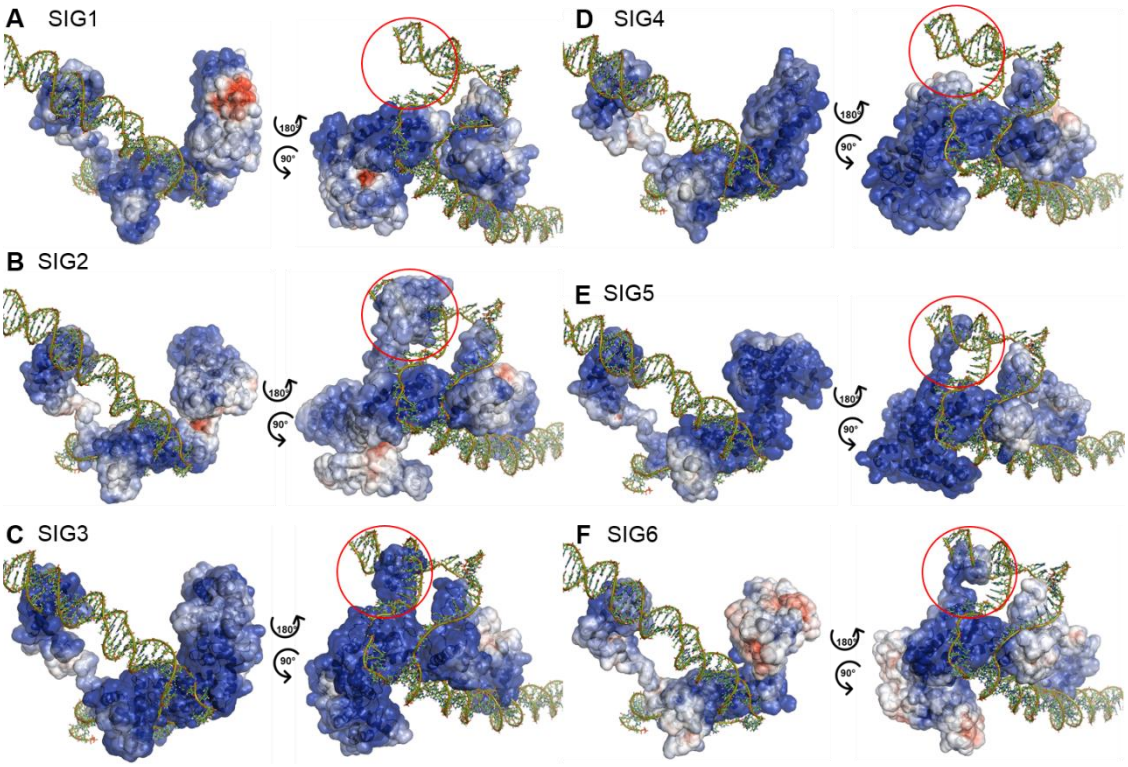


Figure 3-12. DNA-sigma factors models for Arabidopsis sigma factors. DNA from the transcription comlex in *E. coli* (PDB ID 6B6H) was used as template to simulate binding for plant sigma factors A) SIG1, B) SIG2, C)SIG3, D)SIG4, E) SIG5 and F) SIG6. Red circles indicate molecular clashes with DNA only in some sigma factors.

To study how DNA binding might occur within the modelled structures, alignment and sculpting tools were used in PyMOL with the modelled three-dimensional structures of the Arabidopsis sigma factors and a high resolution transcription initiation complex from *E. coli* with DNA bound to it (PDB ID 6B6H). In the template structure we could obtain information regarding the DNA melting step in the transcription initiation complex when rotated 90°.

All sigma factors modelled (Fig. 3-12) have a very positively-charged surface compared to the *E. coli* sigma factor RpoD (Fig. 3-4). All structures bind DNA by their highly positive conserved σ^{70} regions 2 and 4 (Fig. 3-12). However, after examining in more detail the structures rotated 90°, SIG2, SIG3, SIG5 and SIG6 have a protusion that clashed with the DNA (Fig. 3-12). The biggest difference in charge is a negative patch (red) in SIG1 (Fig. 3-12A) which is relatively weak in SIG2 and SIG6 (Fig. 3-12B, 3-12F) but absent in the other structures.

3.2.2 Prediction of sigma factor phosphorylation sites

It has been demonstrated that phosphorylation of T170 in SIG1 (Fig. 3-13A) regulates the balance between components of PSI and PSII and that mutation of an earlier phosphorylation site S55 had no effect in *rbcL* transcript abundance (Shimizu et al., 2010). This raises the possibility that other phosphorylation events might be important regulators of plant sigma factors. To investigate this, experimentally identified and predicted phosphorylation sites were obtained from PhosPhAt 4.0 (Durek et al., 2010) and mapped within the modelled sigma factor structures.

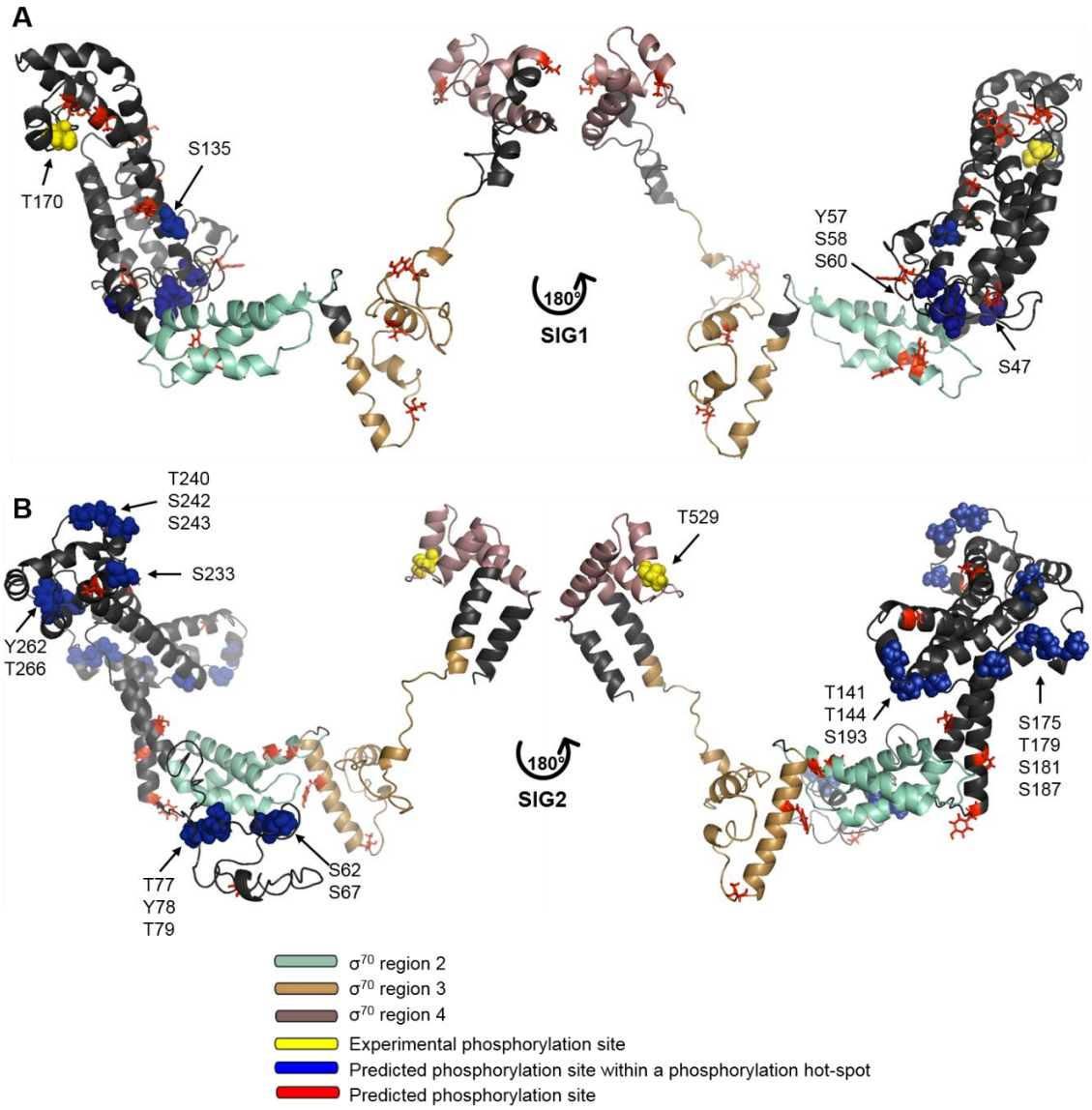


Figure 3-13. Experimental and predicted phosphorylation sites within hotspots are enriched outside the σ^{70} domains in SIG1 and SIG2.

In the SIG2 model, an experimentally-identified phosphorylation site on the σ^{70} region 4 was detected. Since this region is involved in -35 motif recognition, it is likely that the charge and volume from the phosphate group obstruct a possible interaction with the DNA backbone. Therefore, this site is a good candidate for direct mutagenesis to see if the binding targets change in abundance.

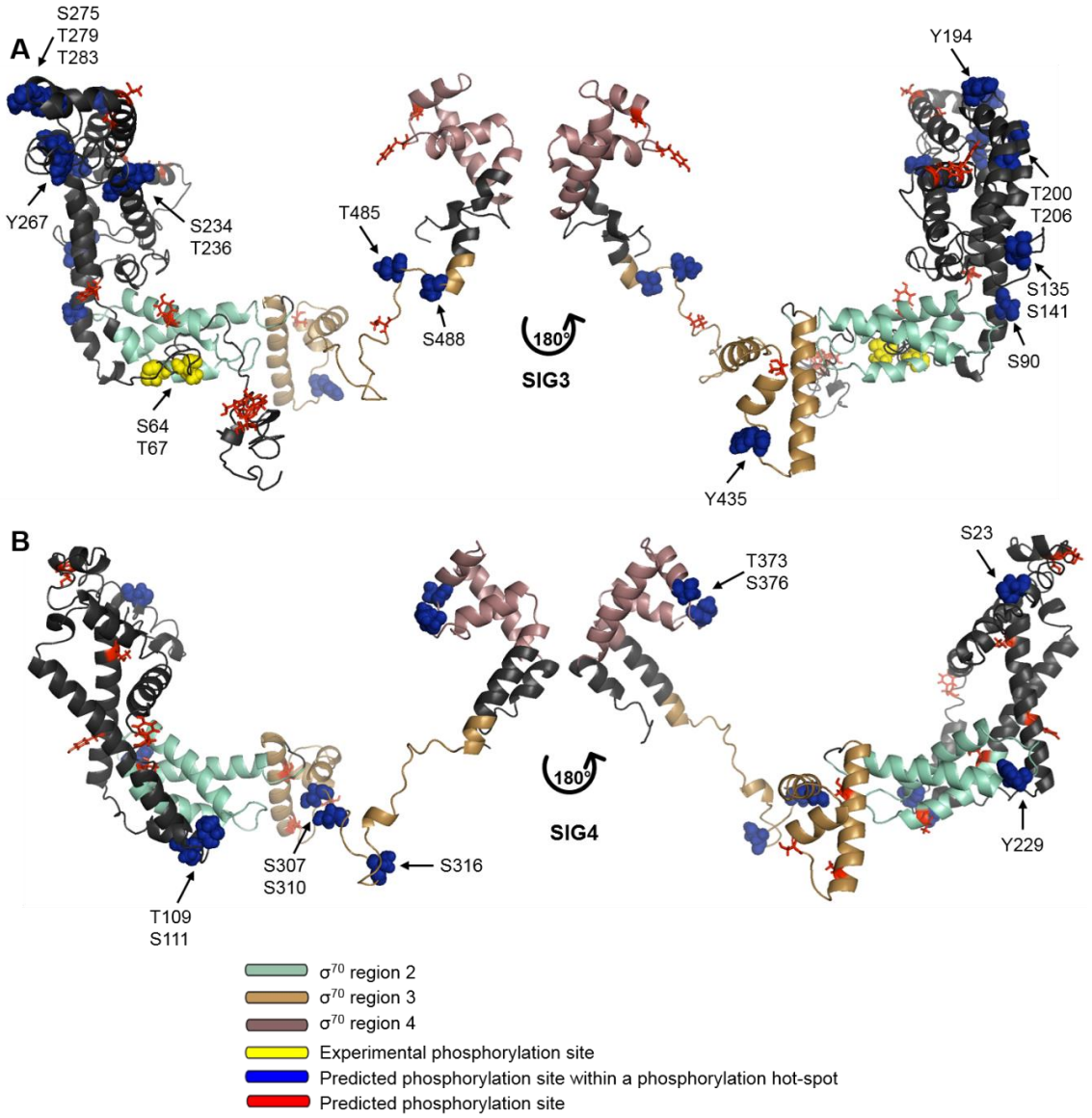


Figure 3-14. Predicted phosphorylation sites within hotspots fall in σ^{70} regions in SIG3 and SIG4.

SIG3 phosphorylation sites mapped to the modelled three-dimensional structure (Fig. 3-14A) indicate that there is an experimentally identified phosphorylation site very early in the protein sequence (N-terminal). The structural model suggests this region is highly disordered, so it is possible that it constitutes part of the signal peptide to chloroplast.

However, there are other phosphorylation sites elsewhere in the protein sequence. The most interesting one is situated in the σ^{70} region 2 in the subdomain that recognises -10 motif sequences.

SIG4 has no experimentally identified phosphorylation sites (Fig. 3-14B), however, as with SIG3 it contains predicted phosphorylation sites within hotspots in the conserved σ^{70} region 2, which could be interesting for further investigation.

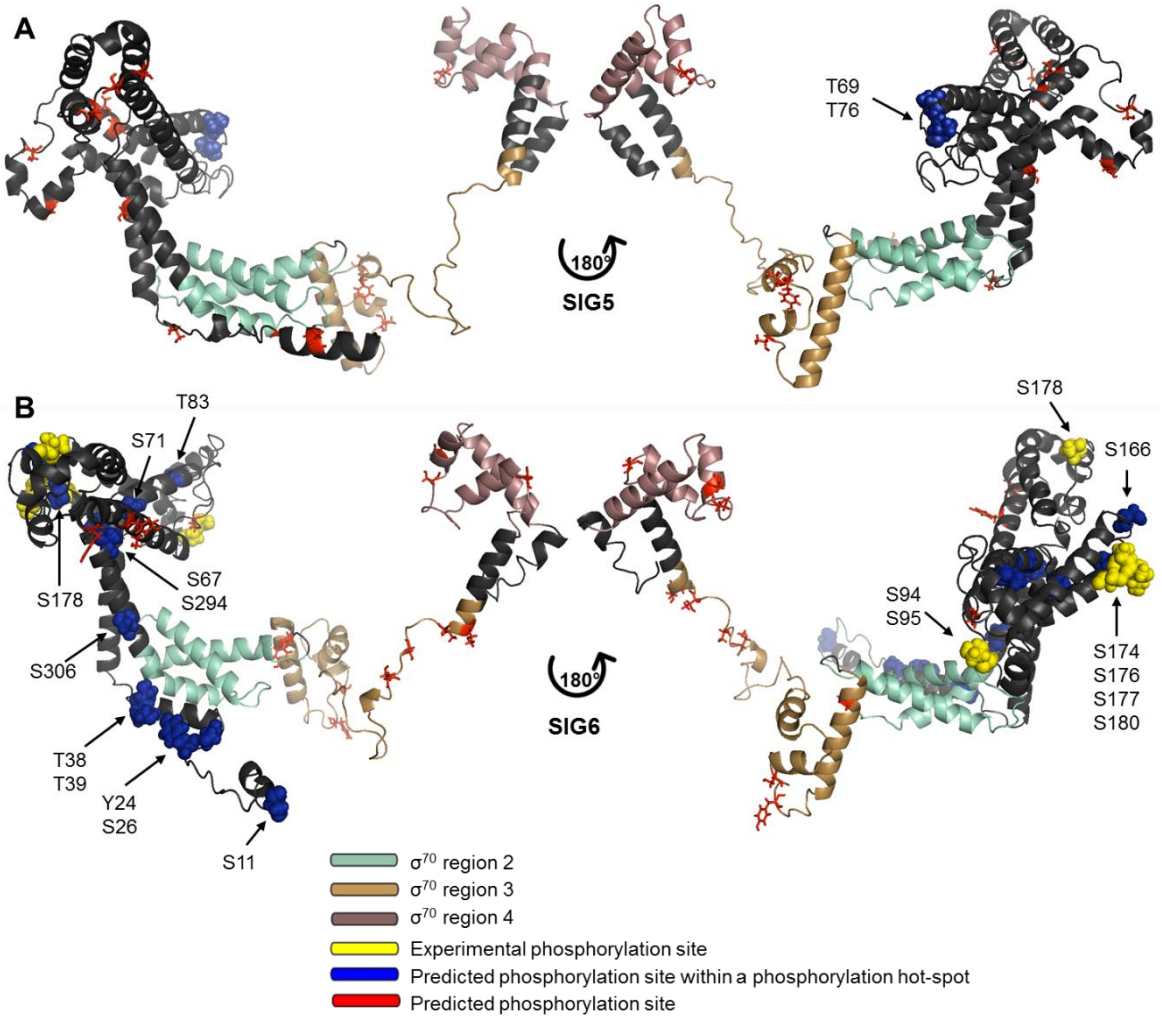


Figure 3-15. There is only one predicted phosphorylation site within hotspots in SIG5 whereas there are plenty of experimentally determined in SIG6.

SIG5 has two predicted phosphorylation sites in a hotspot located in the non-conserved area in a similar position to the sites that regulate SIG1 (Fig. 3-15A). Due to this, we can speculate that T69 and/or T76 might be involved in allosteric changes of promoter recognition in SIG5 as demonstrated for SIG1 (Shimizu et al., 2010).

SIG6 has the greatest number of experimentally identified phosphorylation sites (Fig. 3-15B) (Schweer et al., 2010). The following residues S94, S95, S176, S180 and T249 did not have a phenotypic effect in plants when substituted by alanine, whereas S94/95, S174 and S177 had a phenotype like *sig6* mutants. The authors attributed phosphorylation to a nuclear-coded plastid-targeted casein kinase 2 (cpCK2) (Schweer et al., 2010).

Overall, most of the N-terminal sigma factor segments below the 100 amino acids are enriched in phosphorylation sites, however, no such pattern was found in SIG1 (Fig. 3-13A).

3.3 DISCUSSION

Protein homology modelling has its caveats, one being that the resulting structures are very similar to the protein that was used as the modelling prototype. For example, in this study structures based on the 3IYD have a longer, more rectangular non-conserved region in grey (N-terminal) (Fig. 3-5, 3-7), whereas the structures modelled based on the 5WT1 show a propeller-like arrangement (Fig. 3-6, 3-8, 3-9, 3-10).

The most surprising finding from this study is that plant sigma factors are heavily positively charged, while the *E. coli* main sigma factor RpoD is mostly negatively charged on the surface. Sigma factors in eukaryotes are translated in the cytoplasm, however their activity is in the chloroplast stroma which due to active pumping of H^+ into the thylakoid space during photosynthesis is slightly basic (~ pH 8). It has been shown that the affinity of nuclear DNA binding proteins decreases at higher pH with a two fold difference between pH 7.5 and 8.0 (Blane and Fanucchi, 2015). In this light, the bacterial sigma factor would have lower DNA binding activity in the chloroplast, so I propose that sigma factors in eukaryotes acquired positive surface charges during evolution to increase their DNA binding activity in the chloroplast stroma.

Another possible explanation is that sigma factors incorporate a N-terminal transit peptide required for import into the chloroplast. Transit peptides have little similarity in sequence, which makes them hard to predict, however they are characterised by serine or threonine residues and positively charged amino acids in the N-terminal (Bédard and Jarvis, 2005). Due to this, I think the portions of the sigma factors clashing with the DNA in the putative binding models (Fig. 3-12) are in fact transit peptides, which also explains the highly disordered structures predicted. Supporting this hypothesis are the many

predicted phosphorylation sites in these early segments from the sigma factors. Phosphorylation in the signal peptide to the chloroplast is important for recognition by the receptors in the outer chloroplast membrane and to differentiate them from peptides targeted to the mitochondria which are generally not phosphorylated (Waegemann and Soll, 1996).

Protein import into the chloroplasts is performed by members of the translocon at the outer membrane (Toc) and translocon of the inner membrane (Tic) proteins that form a channel in control of recognising and importing proteins into the chloroplast lumen. In the outer membrane the process requires Toc59 which contains an N-terminal acidic domain that interacts with the positively charged transit peptide, pushing the peptide across the membrane and consuming GTP (Jarvis and Lopez-Juez, 2013). It is possible that the very positively charged sigma factors facilitate this step by diminishing the energy cost of transportation into the intermembrane space. In the inner membrane the process is voltage dependent, however the process is also mediated by a cation channel (Bédard and Jarvis, 2005). Overall, the evolution of plant sigma factors to highly positive proteins suggest that it was favoured by the reduced energy cost of transportation across the chloroplast membranes.

3.3.1 Regulation of plant sigma factors by phosphorylation

Phosphorylation of T170 in SIG1 under conditions of oxidized PQ changes its specificity in promoter recognition reducing the transcript levels of *psaA* (component of PSI) but to a lesser extent the transcription of *psbA* (PSII) (Shimizu et al., 2010). Therefore, phosphorylation of SIG1 is thought to regulate the balance between components of PSI and PSII.

The mechanism by which the phosphorylation site changes promoter specificity in SIG1 is not known yet, however we can speculate it can be caused by either blocking a direct interaction with DNA, supporting the “sliding” mechanism, or due to an allosteric effect where the phosphate makes conformational changes to the whole structure and modifies either the affinity or accessibility of the DNA binding domains.

SIG2 is necessary for transcription of tRNAs, including $tRNA^{Glu}$ which is essential for tetrapyrrole synthesis to produce key components of the light reaction centres in PSI and PSII (Kanamaru et al., 2001). $tRNA^{Glu}$ has also been reported to be involved in the switching between NEP and PEP transcriptional activity by inhibiting NEP activity during

chloroplast development (Hanaoka et al., 2005). $tRNA^{Glu}$ is encoded by the chloroplast gene trne-UUC which contains canonical -35 sequence TTGACA and -10 motif TACTAT in its promoter (Hanaoka et al., 2003). If phosphorylation in the σ^{70} region 4 is weakening the binding of SIG2 to the promoter of $tRNA^{Glu}$ we would expect that SIG2 can be a site of regulation in the chloroplast to either activate NEP transcription, tetrapyrrole synthesis or both during chloroplast development. As with SIG2, SIG6 regulates early chloroplast development and at least two sites have been identified important for its function S174 and S177 (Schweer et al., 2010).

SIG3 is involved in transcription of *psbN* which encodes a protein required for correct assembly of PSII with a special role during photoinhibition (Torabi et al., 2014). The SIG3 N-terminal portion has the highest number of predicted phosphorylation sites structurally close among all N-terminal segments. Photoinhibition can happen in minutes and last hours, therefore rapid and efficient repair must be a priority. It has been shown that phosphorylated chloroplast targeted pre-proteins from wheat are imported 3 to 4 times faster into the chloroplast (May and Soll, 2000), therefore it could be that SIG3 is activated by phosphorylation in response to photoinhibition and rapidly transported into the chloroplast.

SIG4 is responsible for the transcription of a subunit of the plastid NDH complex which is involved in the PSI cyclic electron flow that recycles electrons producing ATP without O_2 breakdown (Shikanai, 2016). This process is especially important when light absorbed is more than the energy used, which means SIG4 is activated during stressful high light conditions.

SIG5 lacks experimentally determined phosphorylation sites and although the predicted site is in the same region as the one modulating SIG1 function, in SIG5 it is very close to the N-terminal portion, which could be due to the phosphorylation site being within the transit peptide.

Depending on their function, it seems that plant sigma factors are likely to be regulated by phosphorylation or by transcription. SIG2 and SIG6 have the greatest number of phosphorylation sites with proven impact on early chloroplast development. This might save energy because the factors do not need to be degraded and then synthesised again when the plants have not developed functional chloroplasts yet, and energy obtained by photosynthesis is not available.

On the other hand, sigma factors involved on photoprotection or response to environmental stimuli have fewer phosphorylation sites. Adult plants are not energy limited and therefore can spend energy resources in stress responses to the environment. Due to this, it could be feasible that established plants degrade and synthesise again proteins that are only used under environmental stresses. However, due to the small time needed for lethal damage to occur in the chloroplasts, sigma factors might have evolved strategies for rapid transportation (SIG3, phosphorylation), contain the minimum domains for function (SIG4) or be circadian regulated (SIG5).

4 SIG5 COMMUNICATES INFORMATION TO THE CHLOROPLAST GENOME CONCERNING LOW TEMPERATURE CONDITIONS

"Cold enough to freeze the balls off a brass(ica) monkey!"

4.1 BACKGROUND

Low temperature is a seasonal, daily or sudden type of abiotic stress that impacts plant growth and therefore restricts geographical distribution and causes losses in food production (Steponkus et al., 1998b). Stress by low temperature (LT) can be divided into two categories: chilling refers to exposure to low non-freezing temperatures, and freezing, which causes cellular dehydration due to exposure to lethal temperatures. A widely studied adaptive response to this environmental stress is cold acclimation. In this, plants exposed to continuous chilling exhibit multiple genetic and metabolic responses that lead to freezing tolerance (Thomashow, 1999).

The best characterized genetic system that responds to chilling and leads to subsequent cold acclimation is the *C-REPEAT BINDING FACTOR (CBF)* cold response pathway (Gilmour et al., 1998;Fowler et al., 2005;Thomashow, 2010). Exposing Arabidopsis to LT induces transcription of *CBF1, CBF2* and *CBF3* (previously known as *DEHYDRATATION RESPONSIVE ELEMENT BINDING1b (DREB1b), 1c* and *1a*). CBFs belong to the AP2/ERF domain family proteins which recognize *C-REPEAT/DEHYDRATION RESPONSE ELEMENTS (CRT/DRE)* contained in the promoters of *COLD REGULATED (COR)* genes such as *COR15A* and *COR78* (also known as *RD29A* and *LT178*) (Thomashow, 2010). Overexpression of these genes confers similar metabolic changes as cold acclimation, but without the need for LT exposure (Gilmour et al., 1998). The role of the circadian oscillator in this pathway has been described previously, *CBF-COR* genes are induced by LT mostly during the day rather than the night, indicating a change in sensitivity to low temperature regulated by the circadian clock that represents circadian gating (Fowler et al., 2005;Dong et al., 2011). This is thought to allow plants to perceive and respond to environmental signalling depending on the time of day, so that adaptive responses are most appropriate for the time of day (Seo and Mas, 2015).

Chloroplasts are vital organelles where photosynthesis takes place. Chloroplasts contain their own genome. The chloroplast genome includes 120–130 genes, which encodes

proteins participating in photosynthesis, transcription, and translation (Daniell et al., 2016). In higher plants, transcription of the chloroplast genome is executed by a nuclear encoded RNA polymerase and a plastid-encoded RNA polymerase, however, nuclear-encoded sigma factors confer promoter specificity to the latter (Tanaka et al., 1997). In Arabidopsis, there are six sigma factors encoded by the nuclear genome (SIGMA FACTOR1 (SIG1)-SIG6) and their expression is circadian regulated (Atkins and Dodd, 2014).

SIGMA FACTOR5 (SIG5) is required for the transcription of the *psbD* blue light-responsive promoter (*psbD* BLRP) in response to blue light and various environmental stresses including high light, high salt, LT and high osmotic conditions (Nagashima et al., 2004). In the Dodd group, it was found that SIG5 communicates timing information from the central circadian oscillator to the chloroplast, where it causes rhythmic transcription of *psbD* and other plastid genes (Noordally et al., 2013).

Exposure of plants to low temperature and high light has been proven to cause photoinhibition, which is a decrease in photosynthetic capacity (Kurepin et al., 2013). Since chloroplasts contain the photosynthetic machinery, support carbon fixation, amino acid synthesis, lipid biosynthesis, among other processes, they must be able to sense when the plant is being exposed to chilling or to lethal low temperatures conditions and when to ignore non-lethal conditions. In this light, we studied the integration of low temperature signalling and circadian timing communication to the chloroplast.

4.2 RESULTS

4.2.1 *SIG5* was induced strongly in response to low temperature

Using eFP Browser online tool available at http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Winter et al., 2007a), I investigated sigma factor induction patterns in response to a treatment comprising cold (continuous 4°C on crushed ice in cold chamber), or hot temperature (at 38 °C and recovery at 25 °C). Only *SIG5* induction changes more than two-fold in response to LT (Fig. 4-1). According to this, three hours or more of continuous low temperature will induce *SIG5* transcript accumulation.

Based on the *in silico* data, under laboratory conditions I investigated which sigma factors were sensitive to LT by exposing WT (Col-0) plants to 19 or 4 °C for three hours and then measured the abundance of all six sigma factors in *A. thaliana* by qRT-PCR. As

expected, only *SIG5* was induced significantly in response to LT treatment in the WT, with an over two-fold increase compared to the control treatment (Fig. 4-2A and 4-2B). This is consistent with previous studies (Nagashima et al., 2004).

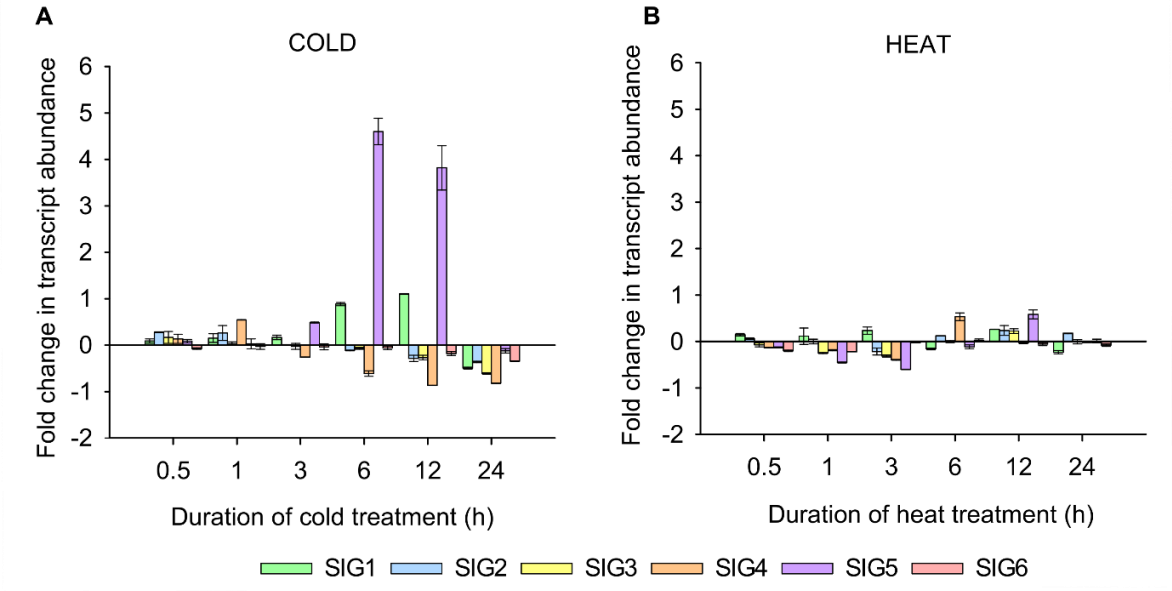


Figure 4-1. Transcriptome data identifies *SIG5* as highly responsive to cold (A) but not high temperature (B). Data indicate the fold-change (4 °C – 19 °C / 19 °C) in abundance of all six Arabidopsis sigma factors during a prolonged chilling treatment (Kilian et al., 2007) extracted using the Arabidopsis eFP Browser (Winter et al., 2007b).

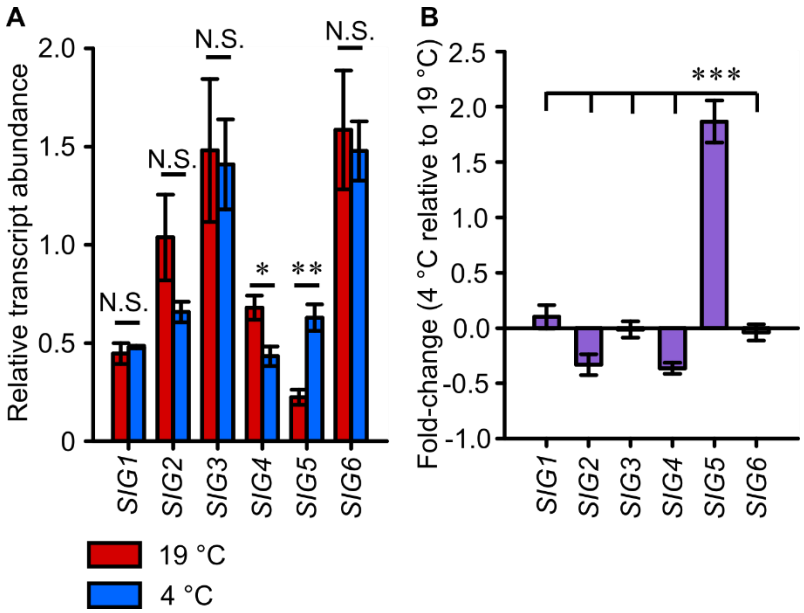


Figure 4-2. *SIG5* is induced strongly in response to LT (A) Relative transcript abundance of all six sigma factors was determined by qRT-PCR in WT (Col-0) plants exposed 3 hours to 19 °C or 4 °C. (B) Fold increase of relative transcript abundance was calculated from values in 1A (4 °C – 19 °C / 19 °C). Data is expressed as means of three independent experiments. Bars indicate s.e.m. (n=3, N≥10 plantlets). (A) shows statistical significance compared to temperature control (4 °C vs 19 °C). (B) shows statistical significance between genotypes (Fold change). N.S., not significant; *P < 0.05, **P < 0.01, ***P < 0.001 (two tailed t-test).

4.2.2 SIG5 communicates information to the chloroplast genome concerning low temperature and this is gated by the circadian oscillator

To test whether the cold-induced increase in *SIG5* transcripts produced a signal that reached the chloroplast and induced the transcription of the *psbD* BLRP, WT and the *sig5-3* loss of function mutant (Noordally et al., 2013) were grown under LD cycles for 11 days, and then exposed to either 19 or 4 °C for 30 minutes, 1, 3, 6, 9, 12 and 24 hours. Relative transcript abundance was then measured by qRT-PCR for *SIG5* and *psbD* BLRP.

LT induced *SIG5* transcript accumulation in WT plants and, as expected, there was no transcript accumulation of *SIG5* in the *sig5-3* plants at any time (Fig. 4-3A). LT also induced *psbD* BLRP transcripts in WT plants 9 h after continuous cold exposure (Fig. 4-3B). However, *psbD* BLRP transcript abundance in *sig5-3* remained low all time. This demonstrated that LT information is being communicated to the chloroplast and this is SIG5 dependent.

It has been demonstrated that SIG5 communicates timing information from the circadian oscillator to chloroplasts where it causes rhythmic transcription of *psbD* (Noordally et al., 2013). Since I have shown that LT is also communicated to the chloroplast by this signalling pathway, I wanted to know whether the circadian timing of the nuclear encoded SIG5 gates the response to LT (whether sensitivity to LT stimulus depends on time of day).

To investigate this, I performed experiments under circadian free running conditions, i.e. plants were under continuous light for 24 h before treatment (LL). Briefly, plants grown under LD cycles were transferred to continuous light and batches of plants exposed to 19 or 4 °C for three hours, starting at different times of day after lights-on (zeitgeber time, ZT).

As expected, *SIG5* transcript abundance was rhythmic (peaks every 24 hours) at 19 °C (Fig. 4-3C) (Noordally et al., 2013). Importantly, the magnitude of *SIG5* induction by LT depended on the time of day when LT pulses were applied to the WT. In this case, a greater response occurred one hour after subjective dawn (ZT 25) compared with later in the day (ZT 37) (Fig. 4-3C). *psbD* BLRP expression was also rhythmic at 19 °C, and the magnitude of induction in response to LT was also gated by the clock, with higher cold-induced transcript abundance 5 hours after subjective dawn (ZT 29) compared to later in the day (ZT 37) in the WT (Fig. 4-3D).

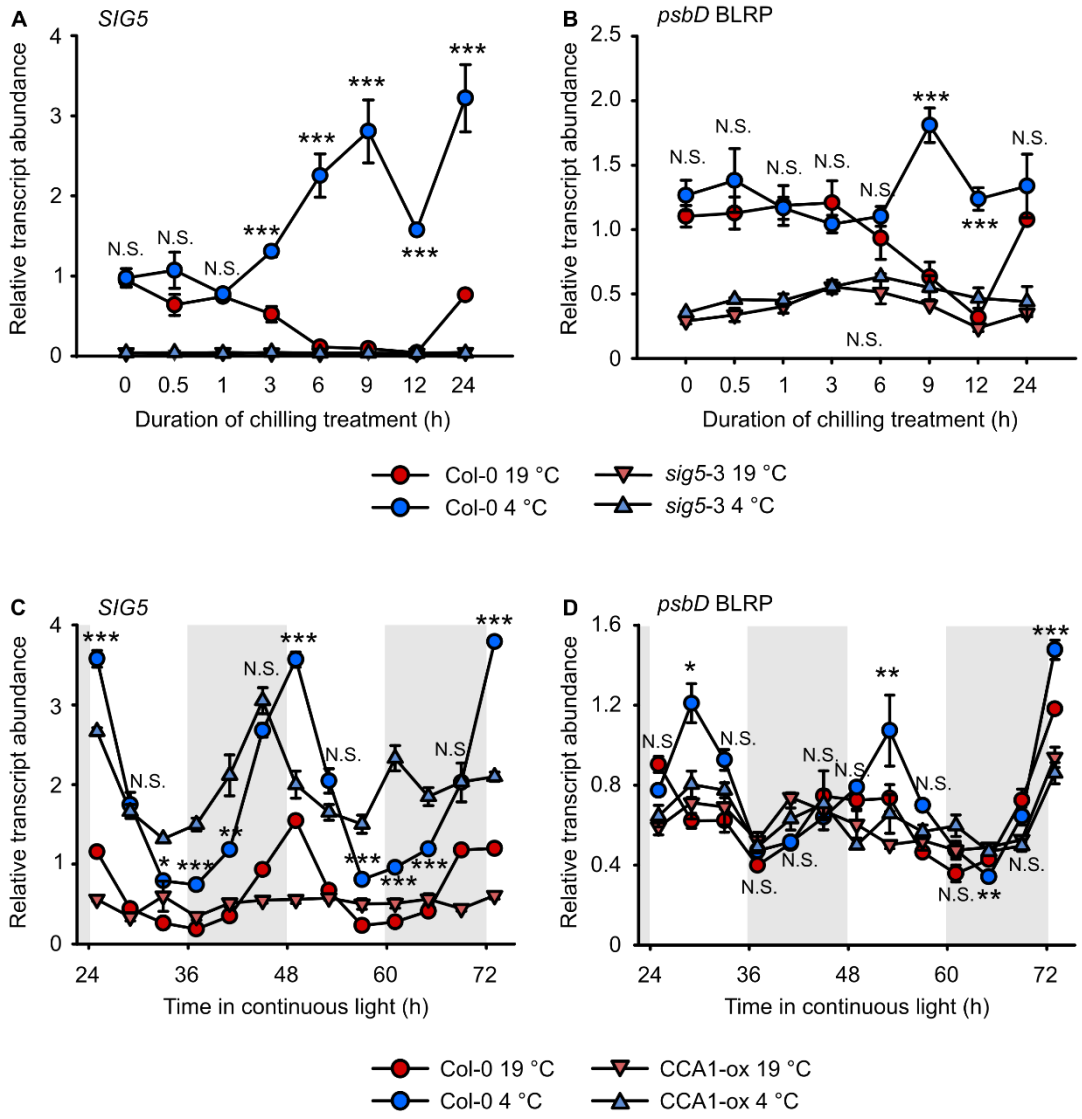


Figure 4-3. The circadian oscillator gates LT signalling to chloroplasts by SIG5 and the response is abolished when the circadian clock is arrhythmic (*CCA1-ox*). Transcript analysis of (A) *SIG5* (B) *psbD* BLRP genes was determined by qRT-PCR in WT and *sig5-3* plants grown at 19 °C for 11 days and exposed to 19 °C or 4 °C for 30 minutes, 1, 3, 6, 9, 12 and 24 hours. (C) *SIG5* (D) *psbD* BLRP, determined by qRT-PCR in WT and *CCA1-ox* plants exposed to 19 °C or 4 °C for three hours starting at the indicated different times of day. Data are expressed as means of three independent biological replicates. Bars indicate s.e.m. (n=3, N≥10 plantlets). (A) and (B) show statistical significance compared to temperature control (4 °C vs 19 °C). (C) and (D) show statistical significance between genotypes at 4 °C (Col-0 4 °C vs CCA1-ox 4 °C). N.S., not significant; *P < 0.05, **P < 0.01, ***P < 0.001 (one-way ANOVA followed by Tukey-Kramer post-test).

If the response of SIG5 to LT is genuinely controlled by the circadian clock, I would expect that this regulation is lost in a transgenic line with a disrupted circadian clock (*CCA1-ox*) (Green et al., 2002). In Fig. 4-3C, *SIG5* transcription was arrhythmic in the

CCA1-ox line at 19°C but at 4°C it exhibited circadian gating of cold induction, however, peaks of the response were shifted, less robust and completed only one 24h cycle under continuous light. Interestingly, when I examined the transcript abundance of *psbD* BLRP (Fig. 4-3D), the induction was abolished comparing *CCA1-ox* and WT at 19 °C but specially compared to WT at 4°C. This result suggests that without a completely functional circadian clock, the increase in *SIG5* transcript abundance by LT did not lead to upregulation of *psbD* BLRP, and instead the expression is similar at different times of day.

In order to assess the role of SIG5 in this gating response to cold, the 3 hour cold pulses were applied to WT and *sig5-3* mutant plants at different times of day under continuous light conditions (Fig. 4-4). LT induced *SIG5* depending on time of day (Fig. 4-4A), which replicated my previous finding that LT activation of *SIG5* is circadian gated (Fig. 4-3C). As expected, there was no accumulation of *SIG5* transcripts in the *sig5-3* mutant plants (Fig. 4-4A).

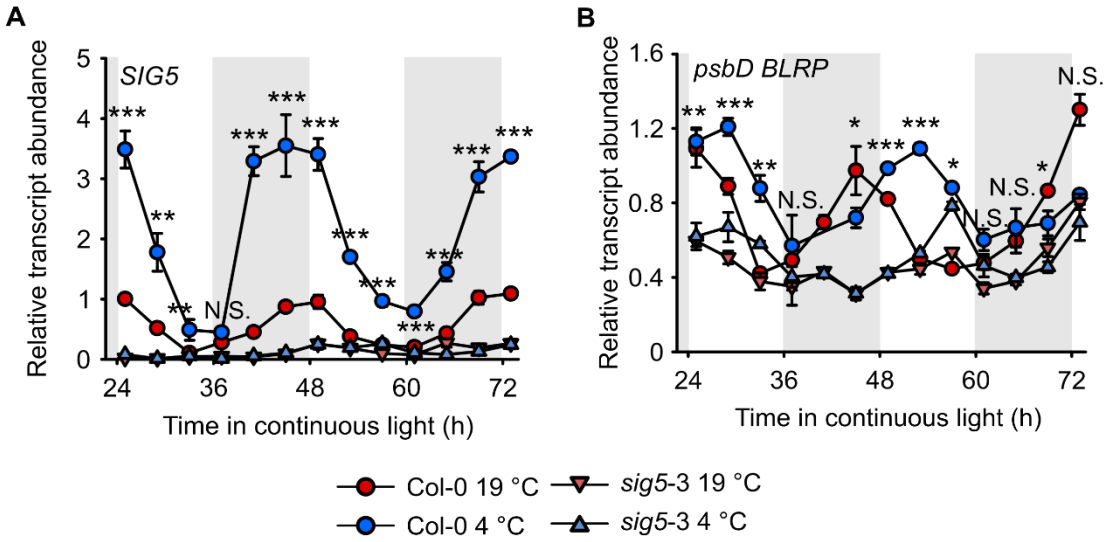


Figure 4-4. SIG5 is required for low temperature transcriptional response in chloroplasts. Transcript abundance analysis of (A) *SIG5* (B) *psbD* BLRP. Each chilling treatment was applied to a separate batch of seedlings. *x* axis indicates time that chilling commenced. Data are expressed as means of three independent biological replicates. Bars indicate s.e.m. (n=3, N≥10 plantlets). (A) shows statistical significance compared to temperature control (4 °C vs 19 °C). (B) shows statistical significance between genotypes at 4 °C (Col-0 4 °C vs *sig5-3* 4 °C). N.S., not significant; *P < 0.05, **P < 0.01, ***P < 0.001 (one-way ANOVA followed by Tukey-Kramer post-test).

Interestingly, *psbD* BLRP in WT plants at 19 °C peaked in abundance four hours before subjective dawn (ZT 45), but when LT pulses were applied the peak shifted to four hours after subjective dawn (ZT 53) (Fig. 4-4B). Circadian expression of *psbD* BLRP at 19 °C and the gating of the response of *psbD* BLRP to LT was absent in the *sig5-3* mutant.

These results suggest that chilling alters the circadian phase of *psbD* BLRP transcript accumulation and this is SIG5-dependent.

4.2.3 HY5/HYH are upstream of the pathway

SIG5 transcript accumulation in response to light is reduced in *hy5* and *cry1* mutants (Nagashima et al., 2004;Belbin et al., 2016). Microarray data suggest that SIG5 is regulated by UV RESISTANCE LOCUS8 (UVR8) (Brown et al., 2005). In this UV-B signalling pathway, the ELONGATED HYPOCOTYL5 (HY5) and HY5 HOMOLOG (HYH) transcription factors mediate UVR8-dependent responses, acting with partial or complete redundancy (Brown and Jenkins, 2008). Also, HY5 has been identified as a LT and light signal integrator because it is degraded by the E3 ubiquitin ligase COP1 in the absence of light at normal temperature, but remains stable in darkness under LT conditions by nuclear depletion of COP1 (Catala et al., 2011).

This led me to hypothesize that *SIG5* transcriptional activation in response to LT could also be mediated by HY5, HYH or COP1. In order to test this, mutants of these genes (*hy5*, *hyh*, and *cop1-4*) and the double mutant *hy5/hyh* were used to assess the induction of five sigma factors in response to LT.

In this experiment, LT had no effect on *SIG1* (Fig. 4-5A) and *SIG3* (Fig. 4-5C), reduced *SIG2* (Fig. 4-5B) and *SIG4* (Fig. 4-5D) expression and induced *SIG5* (Fig. 4-5E) transcripts in the WT plants, (consistent with results obtained in Fig. 4-1A). Interestingly, in the absence of COP1 (i.e. in *cop1-4*), all sigma factors had reduced abundance in response to chilling, possibly due to the pleiotropic role of this protein.

Importantly, *SIG5* was the only sigma factor induced by LT in WT plants. In *hy5* and *hyh* single mutants, *SIG5* was strongly induced in response to LT. However, in the double mutant *hy5/hyh* the induction by LT was absent. Therefore, HY5 and HYH act redundantly, upstream of SIG5, to mediate its induction in response to LT.

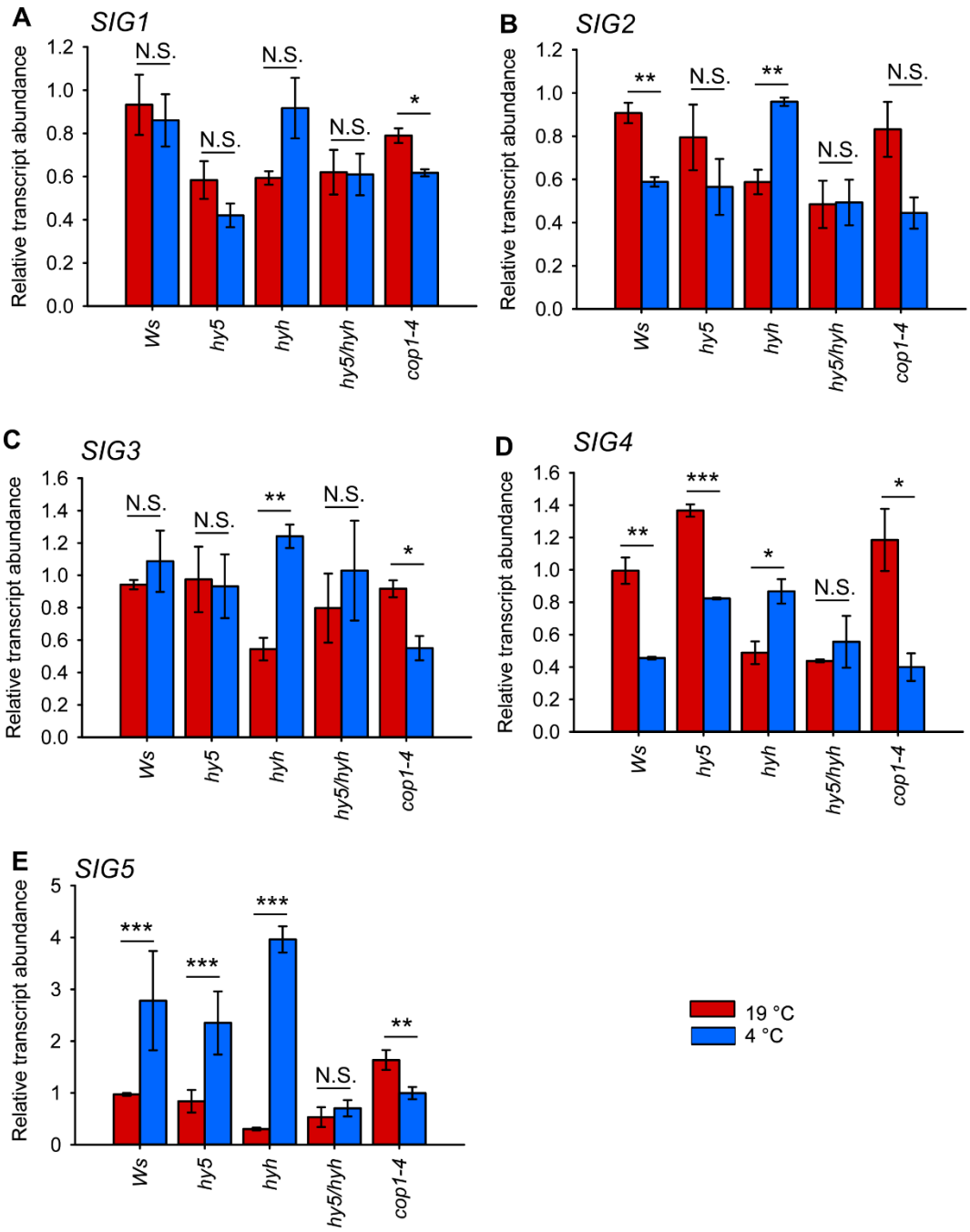


Figure 4-5. HY5 and HYH act redundantly in the cold induction of SIG5. SIGMA FACTORS1-5 transcript abundance was analysed by qRT-PCR in WT, *hy5*, *hyh*, *hy5hyh* and *cop1-4* plants grown under LD cycles for 11 days and then transferred to LL for 24 h. One hour after subjective dawn, plants were exposed to 19 °C or 4 °C for three hours in light. Data is expressed as means of three independent experiments. Bars indicate s.e.m. (n=3, N≥10 plantlets). (A-E) show statistical significance compared to temperature control (4 °C vs 19 °C). N.S., not significant; *P < 0.05, **P < 0.01, ***P < 0.001 (two tailed t-test). Note: *SIG6* was not included in the analysis due to discrepancies in the data using different set of standardised primers.

Since Catala *et al.* (2011) suggest that HY5 might be active in darkness at LT, next I investigated whether light is necessary to induce *SIG5* under cold conditions. To test this, WT, *hy5, hyh*, *hy5/hyh* and *cop1*-4 plants were grown in forward-phase and reverse-phase light/dark cycles for 11 days, then exposed to 19 or 4 °C for three hours in the presence of light (ZT 5, middle of the day for forward-phase plants) or in darkness (ZT 17, middle of the night for reverse-phase plants). I found that *SIG5* can be induced by cold under light or dark conditions in WT plants (Fig. 4-6 A, C). In single mutant plants *hy5* and *hyh* results are similar as in the WT because of the redundancy of HY5 and HYH.

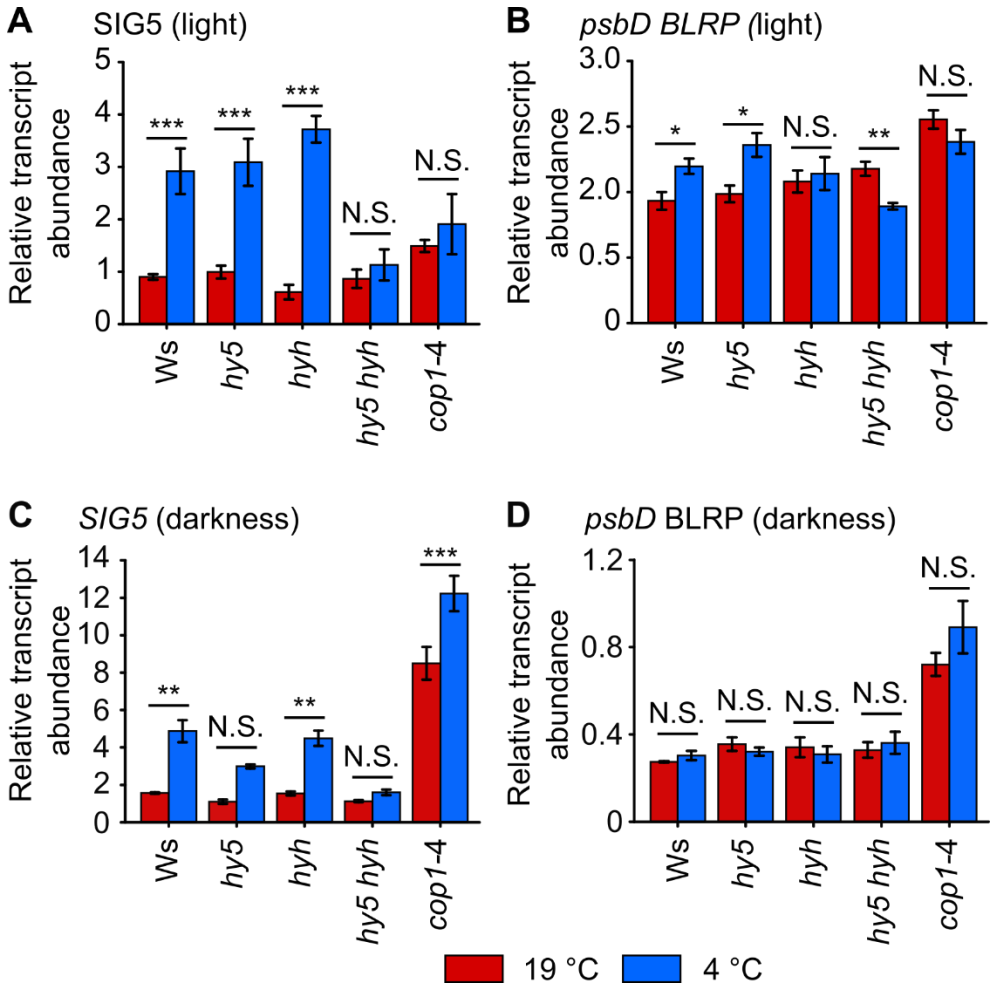


Figure 4-6. Low temperature signalling to chloroplasts by SIG5 is HY5/HYH-dependent and circadian-gated. Responses of (A) *SIG5* and (B) *psbD* BLRP transcripts to chilling in the *hy5 hyh* double mutant and *cop1*-4 mutant in the light. Responses of (C) *SIG5* and (D) *psbD* BLRP transcripts to chilling in the *hy5 hyh* double mutant and *cop1*-4 mutant in darkness (3 hours). Data is expressed as means of three independent experiments. Bars indicate s.e.m. (n=3, N≥10 plantlets). (A-D) show statistical significance compared to temperature control (4 °C vs 19 °C). N.S., not significant; *P < 0.05, **P < 0.01, ***P < 0.001 (two tailed t-test).

Chloroplast *psbD* BLRP was not chilling-responsive in darkness (Fig. 4-6D), potentially due to the light-dependency chloroplast transcription/translation or import to the chloroplast is light dependent. Therefore, SIG5 only communicates information to chloroplasts about chilling temperature conditions in the light (Fig. 4-6B). HY5/HYH are the only mediators of *SIG5* induction at LT in darkness and in constant light, however light is needed to induce chloroplast transcription in response to cold.

4.2.4 SIG5 does not regulate the CBF pathway

As detailed previously, the *C-REPEAT BINDING FACTOR (CBF)* cold response pathway is the best characterized pathway leading to cold acclimation (Fowler et al., 2005) and it has been shown that CBF-dependent genes have a protective role in the chloroplast (Artus et al., 1996). Due to this, I wondered whether there was a connection between both pathways. In order to assess the role of SIG5 in the response of this pathway to chilling, 3-hour cold pulses were applied to WT and *sig5-3* mutant plants at different times of day under continuous light conditions.

CBF1 expression in the *sig5-3* plants peaked at the same times as in the WT (ZT 33 and ZT 57), however, there was an extra peak in transcript abundance 49 hours after dawn in the mutant (Fig. 4-7A). Since there was no clear evidence of effects of SIG5 on *CBF1* expression, I investigated whether other CBF family members or the CBF target gene *COR15a* were altered in *sig5-3* mutant.

Exposing Arabidopsis to LT induces transcription of three closely related proteins CBF1, CBF2 and CBF3 (Gilmour et al., 1998). In *CBF2* transcript accumulation there were no differences in the peaks (ZT 53) between the WT and mutant (Fig. 4-7B). *CBF3* transcription is tightly controlled by the circadian clock and circadian gating of its induction by chilling is very clear (Mockler et al., 2007). Although, the peak response of *CBF3* to LT is at the same time in *sig5-3* as in the WT, the magnitude is either higher or slightly lower compared to the WT when 4 °C was applied 33 and 57 hours after dawn, respectively (Fig. 4-7C). Overall, this suggests that there was no consistent effect of SIG5 upon the *CBF* gene family expression in response to LT pulses.

I also investigated whether SIG5 affects the response to LT of the CBF target *COLD REGULATED15a (COR15a)*. COR15a is targeted to the chloroplast and is involved in membrane remodelling in response to freezing (Steponkus et al., 1998b). *COR15a* transcripts do not appear to be greatly affected in *sig5-3* mutants and no time input is being disturbed since the circadian gating of LT induction of *COR15a* still occurs (gene transcript peaks when stimulus is given in the morning rather than the afternoon) (Fig. 4-7D).

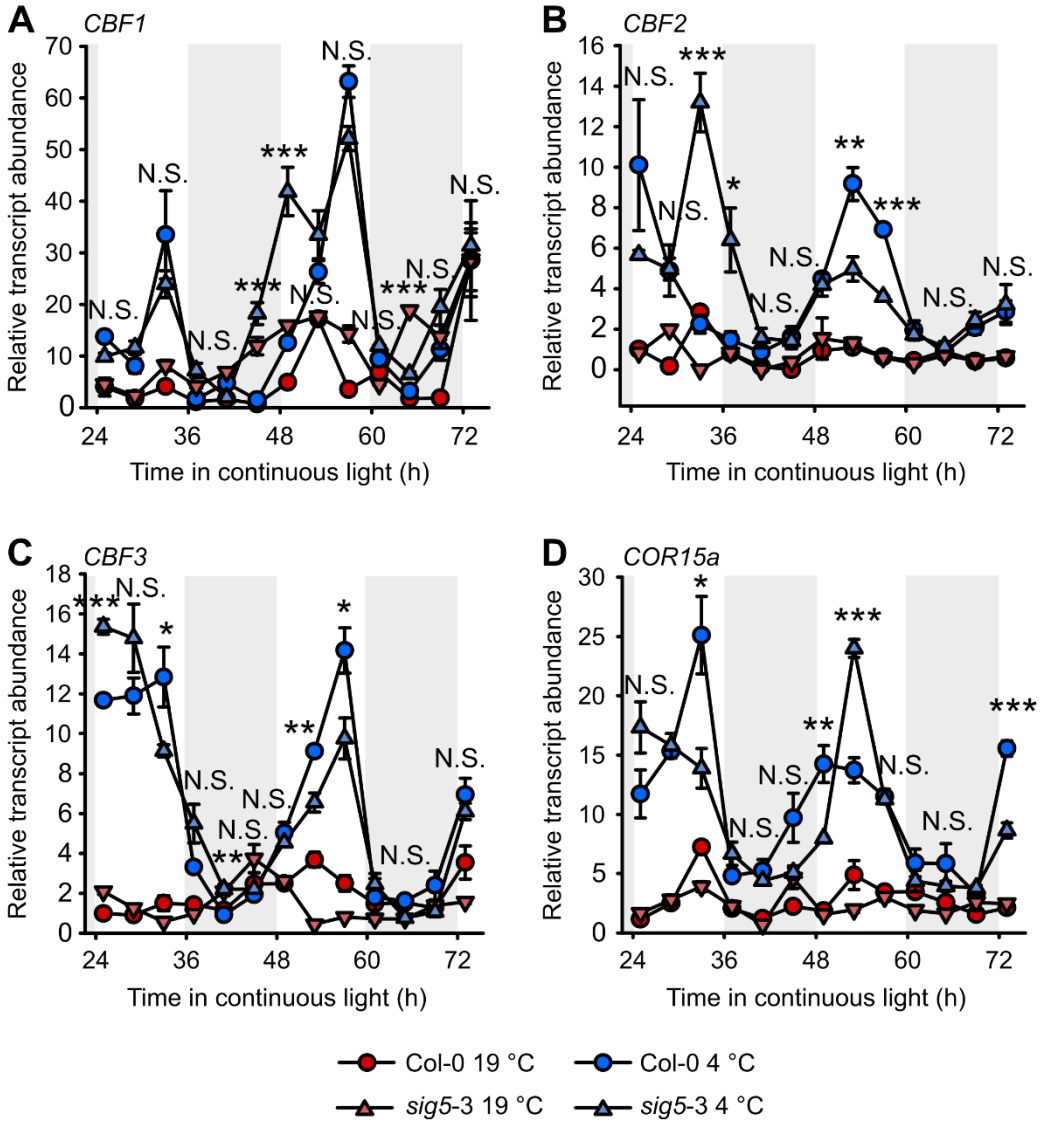


Figure 4-7. Abundance of four transcripts associated with cold adaptation under continuous light. (A) *CBF1*, (B) *CBF2*, (C) *CBF3* and (D) *COR15a* transcript abundance in Col-0 and *sig5*-3 seedlings that were given 3 h chilling treatments commencing at the times indicated on the *x* axis. Data are expressed as means of three independent biological replicates. Bars indicate s.e.m. (n=3, N≥10 plantlets). (A-D) show statistical significance between wild type and mutant specifically at 4 °C (Col-0 4 °C vs *sig5-3* 4 °C). N.S., not significant; *P < 0.05, **P < 0.01, ***P < 0.001 (one-way ANOVA followed by Tukey-Kramer post-test).

4.2.5 *SFR2* is rhythmic, responds to cold and is SIG5-dependent

So far, I have shown that SIG5 is important to communicate temperature conditions to the chloroplasts, however, downstream targets that might lead to environmental adaptation remain unidentified. From previous studies it is known that in addition to CBFs, there are other genes involved in freezing tolerance that modify chloroplast lipid architecture. *SENSITIVE TO FREEZING2* (*SFR2)* encodes a chloroplast targeted protein with beta-glucosidase and galactosyltransferase activity, and mutants in this gene are freezing sensitive (Thorlby et al., 2004;Moellering et al., 2010;Barnes et al., 2016b).

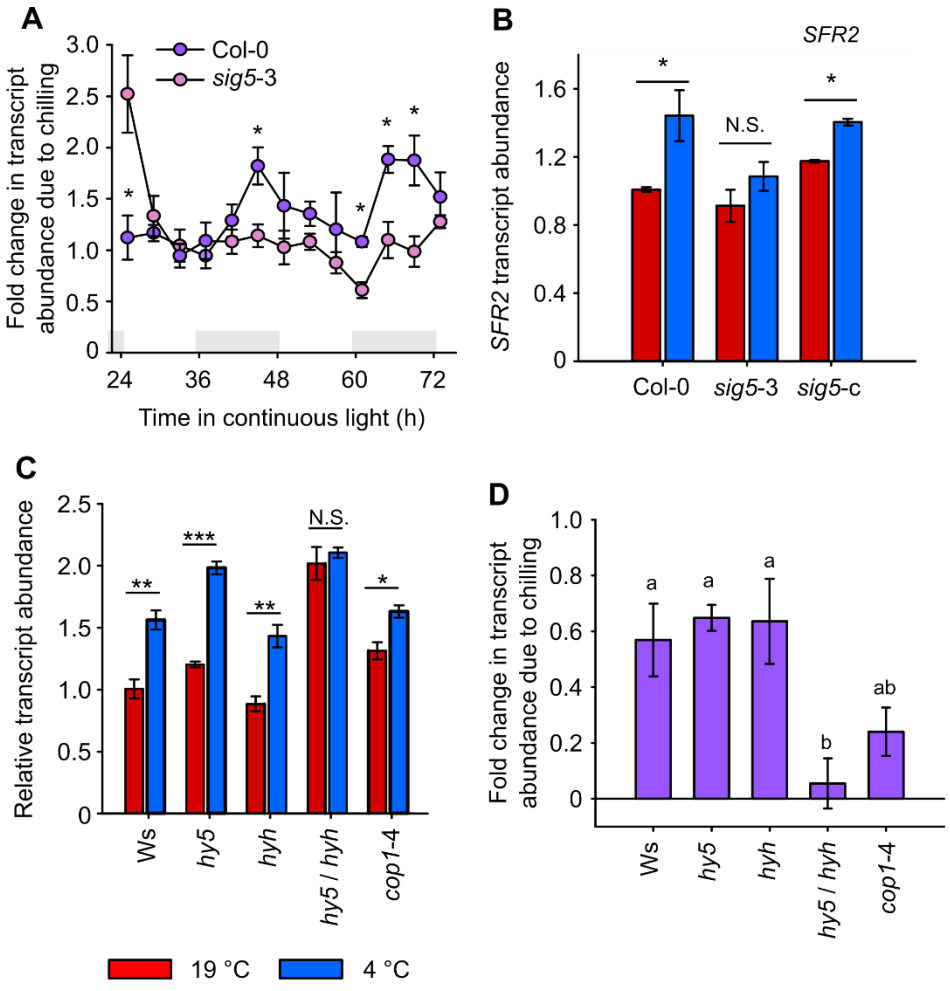


Figure 4-8. SIG5 regulates the accumulation of transcripts encoding the freezing-tolerance protein *SFR2*. (A, B) *SFR2* transcript accumulation and circadian gating in response to chilling requires SIG5. (C, D) Accumulation of *SFR2* transcripts in response to chilling involves HY5, HYH and COP1. (A) was calculated as the difference between each point at 19 °C and 4°C. Bars indicate s.e.m. (n=3, N≥10 plantlets). Analysis by two tailed t-test; N.S., not significant, *P<0.05; ** P<0.01; *** P<0.001, different letters indicate p>0.5.

SFR2 transcripts in WT plants at 19 °C had circadian expression through the day, peaking before subjective dawn and showing lower transcript abundance around dusk (Fig. 4-8A). In response to LT pulses, the level of *SFR2* transcript induction depended on time of day. Surprisingly, *sig5-3* plants had low and arrhythmic expression of *SFR2* at 19 °C, and LT did not induce its transcription (Fig. 4-8B). To verify that the response is SIG5 dependent, SFR2 transcript abundance was determined in a complementation line *sig5-c* (donated by Mitsumasa Hanaoka, Chiba University).

After a three-hour LT pulse at ZT 1 (Fig. 4-8A), *SFR2* was induced in the wild type but not in *sig5-3* plants and the complementation line rescued the response to LT of *SFR2* transcripts. I previously demonstrated that HY5/HYH are upstream regulators of the SIG5 response to LT, hence it was expected that in the *hy5/hyh* double mutant plants, the accumulation of *SFR2* in response to LT would be abolished, if *SFR2* transcript accumulation is regulated by SIG5. *SFR2* transcript was higher in response to LT in the WT, *hy5*, *hyh* and *cop1-4* genotypes, however in the double mutant *hy5/hyh* there was no induction of *SFR2* in response to cold (Fig. 4-8C, D).

4.2.6 A novel SIG5-dependent retrograde signal

Both *SIG5* and *SFR2* are nuclear encoded, however, SIG5 is thought to only regulate transcription in chloroplasts. However, my previous results suggest that SIG5 is upstream of SFR2 induction by LT. To elucidate whether the signal that causes *SFR2* to be induced was coming from the chloroplast (a retrograde signal) I manipulated chloroplast function using inhibitors (Woodson et al., 2013).

These assays consist of applying norflurazon, which causes chloroplast photooxidation (Gray et al., 2003), or lincomycin that inhibits plastid gene expression (Gray et al., 1995). I hypothesised that if *SFR2* induction by chilling needs a functional chloroplast, then in the presence of one of the inhibitors, *SFR2* would be insensitive to LT.

In plants treated with norflurazon, *SIG5* transcript accumulated in response to LT (Fig. 4-9A), however *psbD* BLRP (Fig. 4-9B) and *SFR2* (Fig. 4-9C) transcripts did not accumulate in response to LT. This provides clear evidence of *SFR2* transcription being dependent on a signal from the chloroplast to the nucleus. In the *sig5-3* mutant, regardless of temperature and the presence of norflurazon, no *SIG5* was detected (Fig. 4-9A).

To narrow our search for the SIG5-dependent retrograde signal that acts upon the nuclear encoded *SFR2*, we used a second inhibitor that halts protein translation in the chloroplast. I hypothetised that if SIG5 is causing a chloroplast protein to be transcribed and translated in response to LT that is required for the retrograde signal, the presence of lincomycin would supress this process and *SFR2* would then be unresponsive to LT treatment.

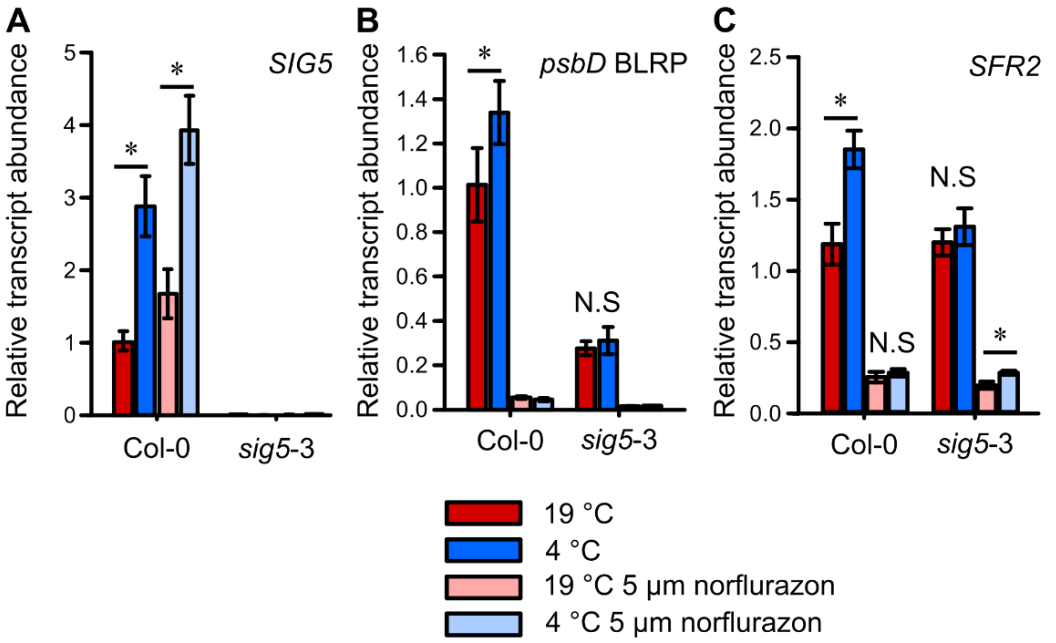


Figure 4-9. Nuclear-encoded chilling responsive genes are regulated by a SIG5-mediated retrograde signal. Accumulation of *SIG5* (A), *psbD* BLRP (B) and *SFR2* (C), transcripts in response to chilling in the presence of 5 µM norflurazon. Data is expressed as means of three independent experiments. Bars indicate s.e.m. (n=3, N≥10 plantlets). (A-C) show statistical significance compared to temperature control. N.S., not significant; *P < 0.05, **P < 0.01, ***P < 0.001 (two tailed t-test).

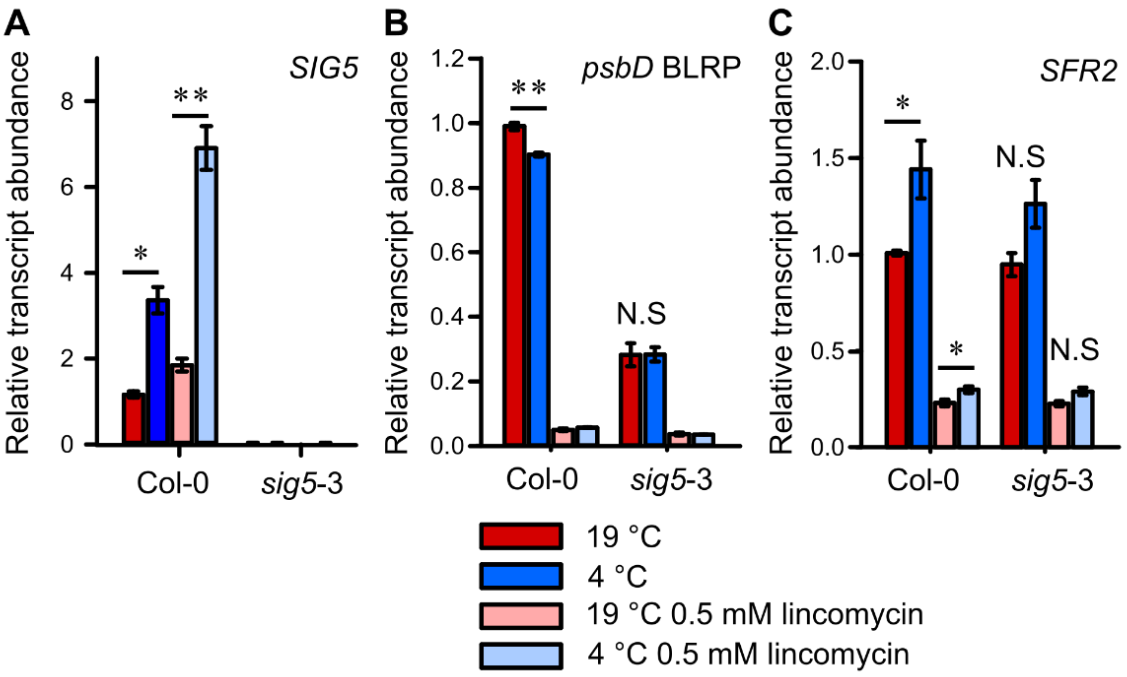


Figure 4-10. Nuclear-encoded chilling responsive genes are regulated by a SIG5-mediated retrograde signal. Accumulation of *SIG5* (A), *psbD* BLRP (B) and *SFR2* (C) transcripts in response to chilling in the presence of 0.5 mM lincomycin. Data is expressed as means of three independent experiments. Bars indicate s.e.m. (n=3, N≥10 plantlets). (A-D) show statistical significance compared to temperature control. N.S., not significant; *P < 0.05, **P < 0.01, ***P < 0.001 (two tailed t-test).

The results show that, as with norflurazon, lincomycin supresses the transcription of *psbD* BLRP and *SFR2* even at 19°C alone (Fig. 4-10). This confirms there is a retrograde signal regulating the transcription of *SFR2* by a SIG5-dependent intermediate translated in the chloroplast.

Because *SIG5* was induced by LT in the presence of both inhibitors, the effect of norflurazon and lincomycin on *SFR2* was specific rather than a genome-wide suppression of transcription.

4.2.7 SIG5 is important to maintain photosynthetic performance (PSII) under low temperature.

So far, we have shown that LT signals are communicated to chloroplasts by a SIG5-dependent pathway, this is gated by the circadian clock, and the response of *SIG5* to cold is mediated by HY5/HYH. I have also shown that SIG5 is required for the activation of a retrograde signal that targets *SFR2* in response to LT. Next, I investigated if this signalling pathway had a direct output on plant performance (photosynthesis, survival) in response to LT or freezing.

Inhibition of PSII activity under strong light is referred as photoinhibition (Murata et al., 2007), a process that has been shown to be accelerated under LT conditions, not by altering the rate of photodamage to PSII but the rate of its repair (Moon et al., 1995). Repair of PSII after photoinhibition involves protein turnover of the photodamaged PSII photosynthetic proteins such as D1 (Takahashi and Murata, 2008) and more recently has been shown to involve protease degradation of D2 in cyanobacteria (Krynicka et al., 2015).

In this LT signalling pathway, SIG5 induces the *psbD* BLRP promoter, which drives the transcription of *psbD*, which in turn encodes the PSII reaction centre protein D2. Since LT information is communicated to the chloroplast via this pathway and induces the transcription of a key component of PSII, and cold accelerates accumulation of damaged photosynthetic proteins (Murata et al., 2007), I wanted to know whether plants that lack SIG5 have altered photosynthetic activity of PSII under cold stress measured as Fv/Fm. The dark adapted values of Fv/Fm reflect the potential quantum efficiency of PSII and are used as an indicator of plant photosynthetic performance (Maxwell and Johnson, 2000).

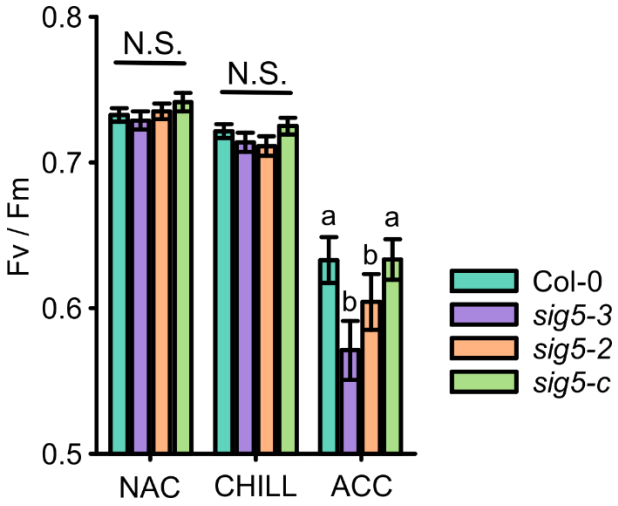


Figure 4-11. SIG5 positively regulates photosynthetic efficiency at LTs. WT, *sig5-3, sig5-2* and *sig5-c* plants grown under LD conditions at 19 °C and then transferred to continous light were exposed to 19 °C or 4 °C during three hours. After the treatment, plants were dark adapted for 15 min before Fv/Fm was calculated. NAC= non-acclimated CHILL= non-acclimated exposed to 4°C for 3 h; ACC = cold acclimated. Data are expressed as means of three independent biological replicates. Bars indicate s.e.m. (n=3, N≥12 plantlets). Statistical significance between genotypes at 4 °C. N.S., not significant; different letters indicate p>0.5 (one-way ANOVA followed by Holm-Sidak test).

To investigate this, WT, *sig5-3*, *sig5-2* and *sig5-c* plants were grown on compost for four weeks under LD cycles and then divided into two groups; non-acclimated and acclimated. Non-acclimated plants remained at 19°C while acclimated plants were moved to 4 °C for 10 days. At the end of the treatment all were transferred to continuous light for 24 h at their respective temperatures. For experiments, a batch of non-acclimated plants was exposed to 4 °C for three hours starting one hour after subjective dawn in constant light (CHILL). After this, non-acclimated (NAC), chilled (CHILL) and acclimated (ACC) were dark adapted for at least 30 min and Fv/Fm was determined.

Analysis of the maximum efficiency of PSII in its dark-adapted state (Fv/Fm) (Fig. 4-11) revealed that WT and *sig5* mutant plants have the same photosynthetic performance at 19 °C, and three hours at 4°C, however, mutation of *sig5* reduces PSII photosynthetic efficiency under long term LT conditions (acclimation), which is rescued in the complementation line. These data demonstrate that besides transmitting temperature information to chloroplasts, SIG5 is important to maintain photosynthetic performance in long term exposure to cold.

4.2.8 Initial damage by freezing assayed through electrolyte leakage

Changes in membrane fluidity during cold acclimation protect plants from freezing damage (Steponkus et al., 1998a;Thomashow, 1999). Electrolyte leakage allows to assess the integrity of the plasma membrane after freezing and thawing and is widely used to quantify freezing tolerance (Hemsley et al., 2014). Cold-acclimated *sig5-3* plants leaked more electrolytes than the WT when frozen to -9.5 °C (Fig. 4-12A). Without cold acclimation, electrolyte leakage was equal across all genotypes at the three temperatures tested (Fig. 4-12B). Results suggest that SIG5 allows long-term chilling to cold acclimate the cells, revealed by the *sig5-3* mutant that suffered damage beyond the chloroplast, impacting plasma membrane integrity at -9.5 °C.

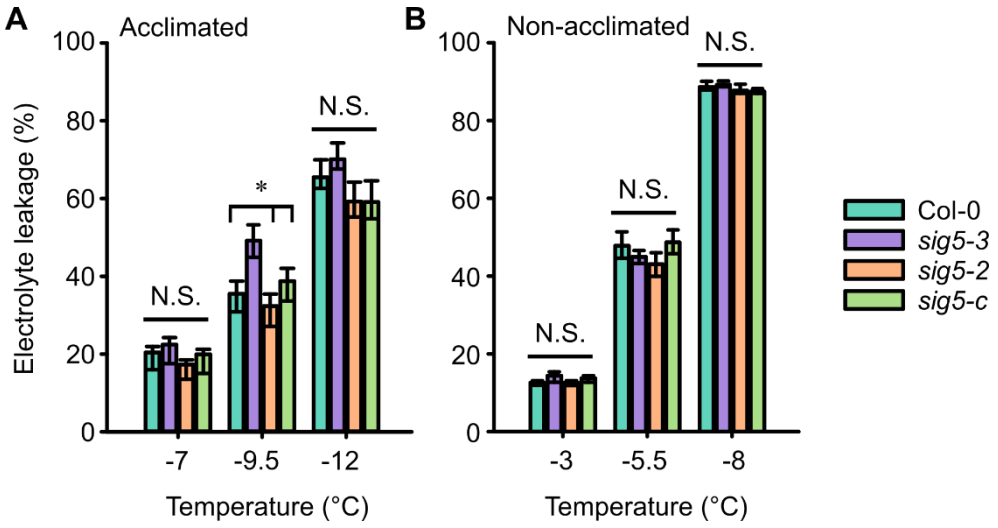


Figure 4-12. Electrolyte leakage from four-week old Col-0, *sig5*-3, *sig5*-2 and *sig5*-c plants that were either cold-acclimated at 5 °C for 14 days or were not cold-acclimated. Data are expressed as means of three independent biological replicates. Bars indicate s.e.m. (n=3, N≥ 6 plants). (A) and (B) show statistical significance between genotypes at each temperature. N.S., not significant; *P < 0.05 (one-way ANOVA followed by Holm-Sidak test).

4.2.9 SIG5 is important for freezing tolerance

Having discovered that SIG5 regulates *SFR2,* a freezing tolerance gene, we wanted to test whether SIG5 is important for survival of whole plants after freezing. WT, *sig5-3*, *sig5-2*, *sig5-c* and *sfr2-3* plants grown for 11 days at 19 °C were then randomly assigned to non-acclimated or acclimated groups and kept for 10 days at either 19 or 4 °C, respectively. On day 21, all plates were exposed to -2 °C for three hours, then ice chips were scattered on top before setting the temperature to -6.0 ± 0.5 °C for non-acclimated

plants and -10.0 ± 0.5 °C for acclimated plants for 24 h. After recovery from freezing at 4 °C for 24 hours and then 19 °C for five days, phenotypes were photographed. After this, plates were dark adapted and after a minimum of 30 min, Fv/Fm was determined and images were recorded.

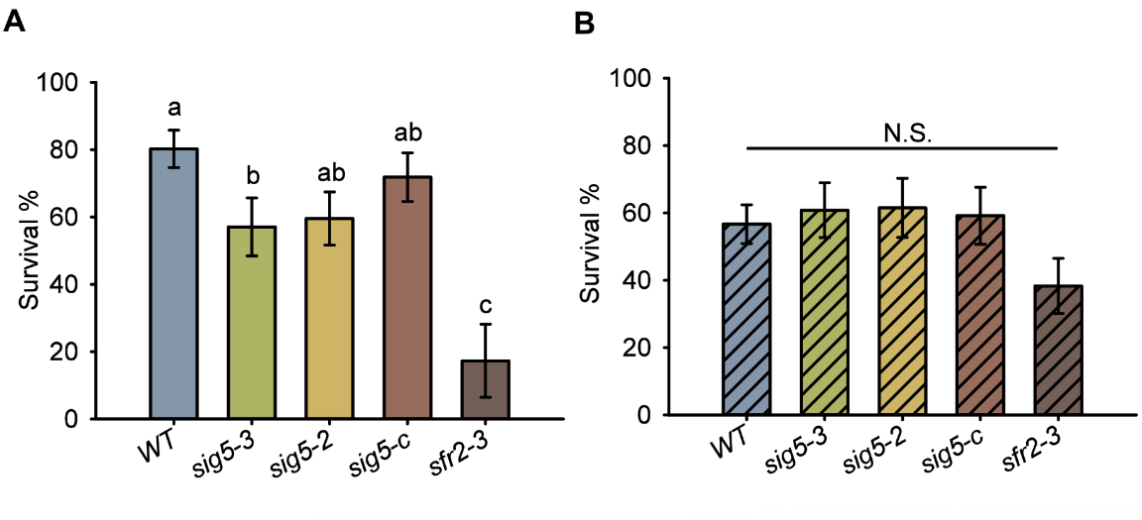


Figure 4-13. SIG5 is required for freezing tolerance. Non-acclimated (A) and acclimated (B) 21-day old WT, *sig5-3*, *sig5-2, sig5-c* and *sfr2-3* plants were exposed to -2 °C for three hours, then ice nucleated and kept to -6 °C or 10 °C for 24 h respectively. Plants were returned to 19 °C for five days for recovery followed by dark-adapted state Fv/Fm determination. Survival was assessed by phenotype and photosynthetic efficiency and %survival calculated. Data are expressed as means of six independent biological replicates. Bars indicate s.e.m. (n=6, N≥ 6 plants). (A) and (B) show statistical significance between genotypes at each temperature. N.S., not significant; different letters indicate P>0.5 (one-way ANOVA followed by Student-Newman-Keuls Method post-test).

The analysis of the images was performed blind, with the help of the Ph. D student Paige E. Panter (Durham University). The plants were recorded as alive only when 1) there was a green apical meristem and 2) in the Fv/Fm coloured scale image, plants appeared green, blue or purple (values >0.6 of Fv/Fm). Survival was calculated as alive plants over total plantlets (%). Fig. 4-13A shows that non-acclimated WT, *sig5-2* and *sig5-c* plants resist -6 °C but *sig5-3* and *srf2-3* had less than 50% survival, which means that they were severely affected by freezing. In acclimated plants there was no difference across all genotypes (Fig. 4-13B). Therefore, SIG5 is not essential for survival in acclimated plants but it is for freezing tolerance.

4.2.10 Lipid remodelling in chloroplasts as an output of the pathway.

Death by freezing in plants is not caused by freezing, but instead by dehydration at the time of thawing (Steponkus and Lynch, 1989a). Ice forms outside the cell which diminishes the water potential causing its exit from the cells. At this point, volume decreases, and membranes come into contact, for example the plasma membrane and the chloroplast outer membrane. Depending on their lipid composition, some of the membranes might fuse together. Non-bilayer hexagonal II (H_{II}) structures are more likely to cause membrane fusion due to abundance of lipids with a conical shape (smaller hydrophilic moieties), whereas lipids in cylindrical shapes are less prone to cause membrane fusion (Steponkus and Lynch, 1989b). During thawing, water enters into the cell, increasing the intracellular volume and if membranes fuse, this causes rupture of the cell and death.

SFR2 activity produces oligogalactolipids by transferring the galactose moieties from one galactosyldiacylglycerol to another, resulting a oligogalactosyldiacylglycerol (cylindrical shape) and the release of a diacylglycerol molecule (Moellering et al., 2010;Barnes et al., 2016a). SFR2 activity results in the production of diacylglycerol, which eventually is transformed to triacylglycerol (TAG) and accumulated in the TAG pool. The TAG product of SFR2 activity is characterised by the presence of 16:3 and 18:3 acyl groups (Moellering et al., 2010;Barnes et al., 2016a). Because SIG5 regulates *SFR2* transcription, we tested whether SIG5 influences lipid remodelling by measuring the changes in TAG content and composition during freezing of cold-acclimated or non-acclimated plants.

The lipid analysis was performed by Tokiaki Takamura under the supervision of Dr. Sousuke Imamura at the Tokyo Institute of Technology, Japan.

Freezing caused TAG accumulation in WT plants, regardless of previous acclimation (Fig. 4-14A). Therefore, I will focus on the comparison between non-acclimated plants. Freezing increased TAG accumulation in the WT and *sig5-c* line, whereas *sfr2*-3, *sig5*-3 and *sig5*-2 accumulated at least 40% less TAG than the WT (Fig. 4-14A). Furthermore, the proportion of 16:3 and 18:3 acyl groups within TAG increased in the WT and SIG5-c plants during freezing but not in *sig5-3* and *sig5-2* mutants (Fig. 4-14B), as in the *sfr2*-3 mutant (Moellering et al., 2010).

Total lipid analysis (Fig. 4-15) revealed that *sig5* mutants had a different TAG fatty acid profile before any treatment. The TAG pool in the mutant plants was characterised by enrichment with 16:1 and 16:3 acyl chains at 19 °C, but these were absent from the WT, *sfr2*-3, and *sig5-c*. Surprisingly, we found that these species were only detectable after freezing in the WT. Therefore, SIG5 appears to influence lipid composition through mechanisms including and additional to the regulation of *SFR2* transcription.

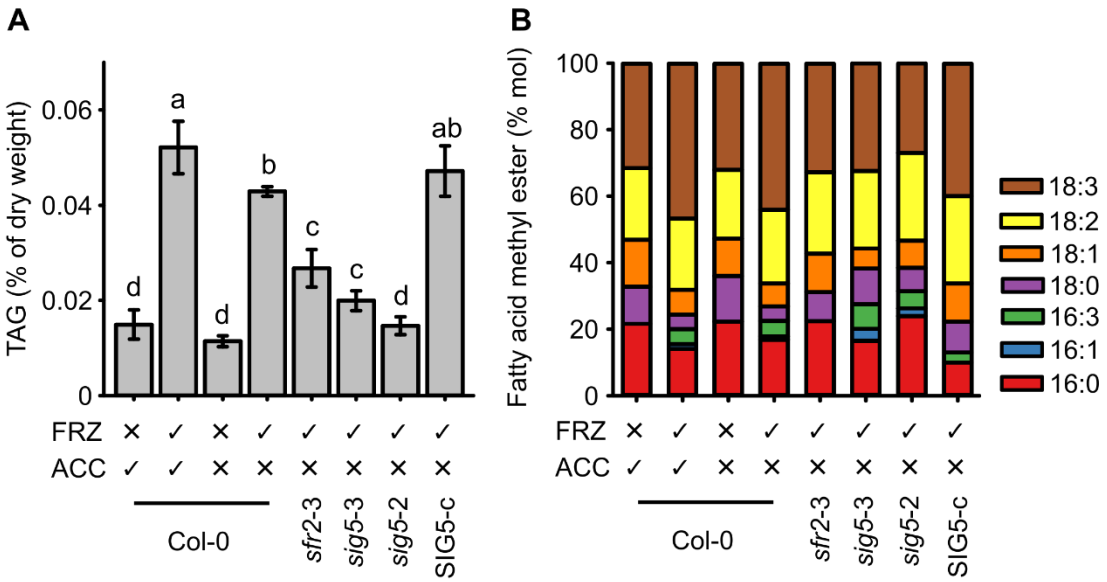


Figure 4-14. Alterations in (A) total triacylglycerol content (TAG) and (B) fatty acid composition of TAG of WT, *sfr2*-3, *sig5*-3, *sig5*-2 and *sig5*-c seedlings. Freezing treatments included three hours at -2 °C, ice nucleation and 24 h at -6 °C. FRZ= freezing and ACC = cold acclimation. Data is expressed as means of three independent experiments. Bars indicate standard deviation and different letters above them indicate significant differences between samples p>0.5. Data was analysed by ANOVA and post-hoc Holm-Sidak test. The shortest descriptions of fatty acids include only the number of carbon atoms and double bonds in them (e.g. C18:0 or 18:0). C18:0 means that the carbon chain of the fatty acid consists of 18 carbon atoms and there are no (zero) double bonds in it, whereas C18:1 describes an 18-carbon chain with one double bond in it. This nomenclature is used within this figure.

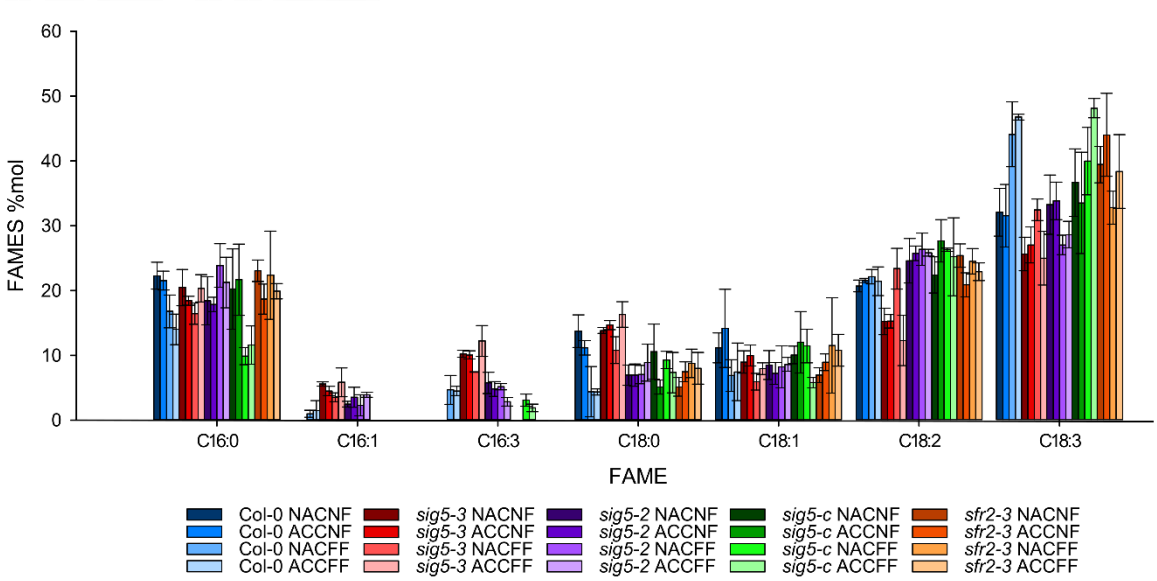


Figure 4-15. Fatty acid methyl ester (FAME) composition of TAG within WT, *sig5-3, sig5-2, sig5-c* and *sfr2-3* seedlings. Quantity of each FAME within TAG of cold-acclimated seedlings that were not frozen (ACCNF) or frozen (ACCFF), and non-acclimated seedlings that were not frozen (NACNF) or frozen (NACFF). Treatments: ACC = Acclimated to 4°C for 10 days and not frozen. NAC = not cold-acclimated. FF = Subjected to freezing for 24 h at -6 °C. NF = Not frozen. The shortest descriptions of fatty acids include only the number of carbon atoms and double bonds in them (e.g. C18:0 or 18:0). C18:0 means that the carbon chain of the fatty acid consists of 18 carbon atoms and there are no (zero) double bonds in it, whereas C18:1 describes an 18-carbon chain with one double bond in it. This nomenclature is used within this figure.

4.3 DISCUSSION

4.3.1 On the identity of the retrograde signal

Plants are exposed to very variable environmental conditions throughout the days and seasons and must adjust their photosynthetic capacity accordingly. For example, it has been shown that under high light stress plants can adjust the number of light harvesting antennas in the chloroplast to avoid damage (Bräutigam et al., 2009). Exposure of plants to low temperature and high light has been proven to cause photoinhibition, which is a decrease in photosynthetic capacity when photon flux is excessive and protein turnover slows down due to low temperature (Kurepin et al., 2013). Because many processes that lead to freezing tolerance occur in the chloroplasts, an efficient communication system must exist between the nucleus and the plastid.

PQ/PQH_2 redox status can also be altered in low temperature conditions (Kurepin et al., 2013). PQH_2 imbalance induced by DBMIB-DCMU treatment has been shown to upregulate *CBF1* and *CBF3* (Bode et al., 2016). Additionally, it has been demonstrated that the PQ redox status is fine-tuned through phosphorylation of SIG1 that regulates its activity and modulate the ratio between PSI And PSII (Shimizu et al., 2010).

In this study we have shown that there is a two-way communication system between the nucleus and the chloroplasts that integrates time of day and low temperature signals to activate adaptive mechanisms in photosynthesis performance and chloroplast membrane remodelling. The key finding is that nuclear-encoded genes respond to the functional state of the plastid under LT stress.

An interesting finding is that in the presence of norflurazon and lincomycin, *SIG5* basal levels are higher than in the WT without the inhibitors and the induction in response to LT is greater (Fig. 4-9, 4-10). This suggests that SIG5 is also part of a retrograde signal in response to a redox imbalance under LT and there is a demand for components of the photosynthetic machinery to restore the redox homeostasis. In the case that the redox imbalance due to cold has a unique signature, then freezing tolerance genes might be evoked alongside SIG5. However, we can not ignore the evidence that *SFR2* induction requires SIG5, which does not fit in this possible explanation.

Therefore, the biggest unanswered question is whether the changes in nuclear transcription are truly SIG5-dependent and signal specific or are indirect effects of a systematic energy imbalance due to low temperature.

Analysis with ARAPORT (v1.10.4, Araport11; available at https://www.araport.org/ accessed 11/09/2018) identified a set of genes that co-express with *SIG5* (39 genes) and the terpenoid backbone biosynthesis pathway was found significantly enriched (p=0.044):

Table 4-1. Pathway enriched components in the co expression gene list retrieved from Araport for *SIG5*.

Identifier	Name	Description
AT5G58770	cPT4	Undecaprenyl pyrophosphate synthetase family protein
AT1G78510	SPS1	Solanesyl diphosphate synthase 1
AT1G17050	SPS2	Solanesyl diphosphate synthase 2

Genes in Table 4-1 are involved in plastoquinone synthesis by condensation of isopentenyl pyrophosphate (IPP) and geranylgeranyl diphosphate (GGPP). IPP can come from two biosynthetic pathways; the mevalonate or the methylerythritol phosphate pathway (MEP). However, it was found that IPP from the MEP pathway is used in the chloroplast (Lange et al., 2000). The backbone molecule for IPP is HMBPP which comes from MEcPP. From the literature we know that MEcPP is a retrograde signal molecule accumulated in response to wound and high light stress (Xiao et al., 2012).

The MEcPP pathway was discovered in the mutant line *ceh1* that lacks 1-hydroxy-2-methyl-2-(E)-butenyl4-diphosphate synthase (HDS), and consequently accumulates high levels of MEcPP which in turn increases the production of salicylic acid (SA) (Xiao et al., 2012). Recently, by transcriptomic, proteomic and metabolomic approaches the MEcPP and SA signalling pathways were dissected using a double mutant defective in ISOCHORISMATE SYNTHASE1 (*eds16*) that is incapable of accumulating SA in the *ceh1* background (Bjornson et al., 2017).

We found *SIG5* and *SFR2* transcripts in the differentially expressed gene list from that study, and we obtained the following results:

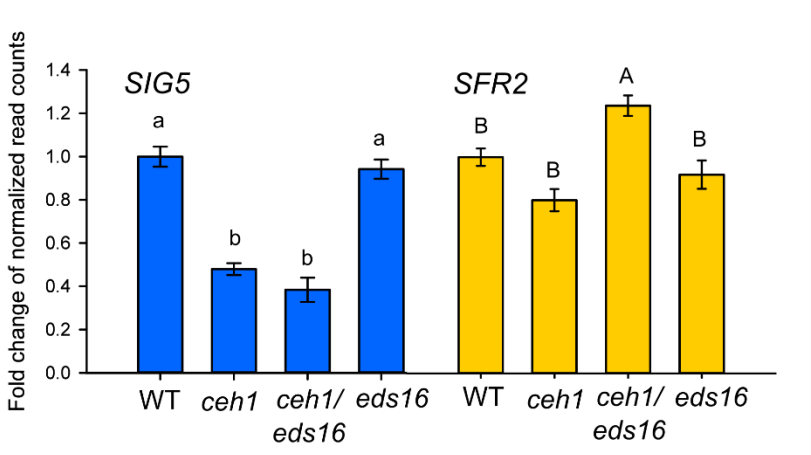


Figure 4-16. Accumulation of MEcPP but not SA induces SFR2 transcript accumulation. Data retrieved from (Bjornson et al., 2017). Different letters indicate significant differences. ANOVA with Holm-Sidak test.

In the line accumulating MEcPP and SA (*ceh1*) and the line that accumulates only MEcPP (*ceh1/eds16*), *SIG5* is repressed compared to the WT and single mutant *eds16* (Fig. 4-16, 4-17A). *SFR2* is induced only when MEcPP accumulates but not SA (Fig. 4-16, 4-17B). This suggests that MEcPP is involved in the retrograde signal that induces *SFR2* but not *SIG5* and SA has an inhibitory effect on both transcripts (Fig. 4-17). Therefore, even if *SIG5* is part of a retrograde signal, it is not the same inducing *SFR2*.

We then can hypothesise that SIG5 communicates low temperature and timing information to the chloroplast and this is required for MEcPP retrograde signal to induce the SIG5-coexpressed genes *Solanesyl diphosphate synthase 1/2* (*SPS1/2)* (Fig. 4-17C). Because SPS1/2 once in the chloroplast uses the carbon backbones from the MEP pathway, MEcPP accumulation could be relieved through the following enzymatic reactions to produce plastoquinone, leaving no time for SA accumulation. The *ceh1/eds16* double mutant has higher levels of DMAPP than *ceh1,* which is two enzymatic steps after MEcPP (Bjornson et al., 2017), supporting the hypothesis that avoiding the depletion of MEcPP in the plastid due to signalling to the nucleus yields higher levels of more complex metabolites further in the MEP pathway (Fig. 4-17C).

In the case of the *sig5-3* mutant, without SIG5 signalling to the chloroplast no SPS1/2 is induced, causing MEcPP sustained accumulation and in consequence that the plastoquinone pool is not replenished leading to ROS production (Fig. 4-17D). Finally, an interplay between the redox balance, metabolite and hormones could be involved in the chloroplast lipid remodelling in response to cold.

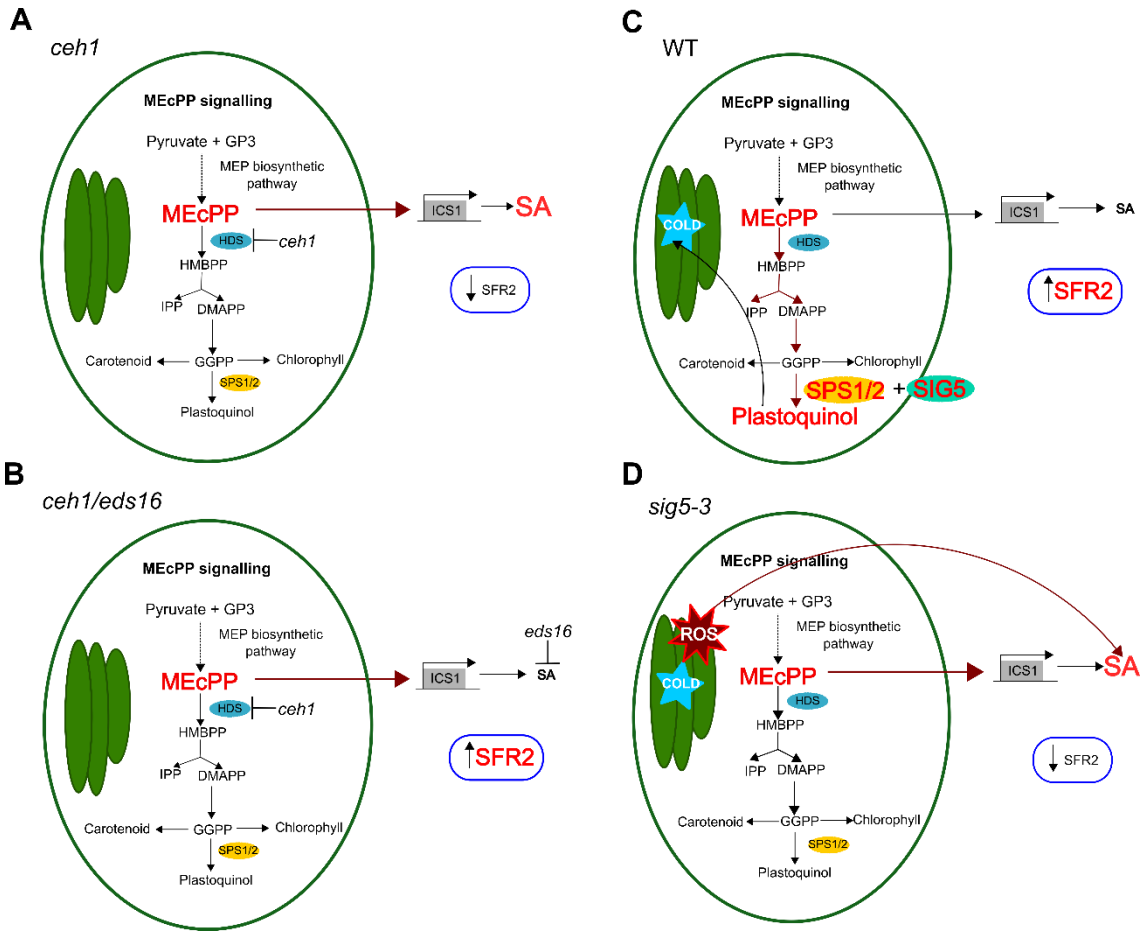


Figure 4-17. Schematic representation of the model proposed for the SIG5-dependent retrograde signal mechanism in response to low temperature. Model showing the MEcPP signalling pathway under stress in the *ceh1* (A), *ceh1/eds16* (B), WT (C) and *sig5-3* (D). Red arrows show active processes and red letters indicate accumulation.

To explore this in future, I propose to use DBMIB to mimic PQ reduced state under low temperature (all components are reduced) in WT plants and measure *SIG5*, *SFR2*, psbD BLRP transcript abundance at 19 °C under LD conditions. Also, MEcPP and SA levels would be measured.

4.3.2 SIG5 in the regulation of eukaryotic lipid biosynthesis.

A surprisingly result from the lipid composition experiments was that *sig5* mutants have TAG with 16:1 and 16:3 acyl chains which are absent in WT plants at 19 °C. Many land plants incorporate prokaryotic and eukaryotic pathways for plastid lipid synthesis. These are differentiated by the length and saturation of the fatty acids incorporated into lipids, with 16:1/16:2/16:3 and 18:2/18:3 fatty acids produced predominantly by the prokaryotic and eukaryotic pathways, respectively (Benning, 2009). The presence of 16:1/16:3 fatty acids in TAG from *sig5-3* but not WT plants at 19 °C (Fig.4-15) suggests that SIG5 might contribute to the balance of lipid synthesis between these pathways. Both SIG5 and the diversification plastid lipid synthesis pathways apparently emerged during the colonization of the land (Kanazawa et al., 2013;Mongrand et al., 1998), so these mechanisms might have related evolutionary histories. Furthermore, rice is chilling sensitive and is a 18:3 plant (Zheng et al., 2016), it would be interesting to see whether *SIG5* overexpression in rice increases the diversity of lipids and ultimately it makes the crop chilling resistant.

5 SIG5-MEDIATED SIGNALLING TO CHLOROPLASTS UNDER NATURAL CONDITIONS

"2+2=5"

5.1 BACKGROUND

The coordination of biological processes with daily and seasonal changes in the environment is important for the survival of plants (Dodd et al., 2005). To achieve this, plants must extract information from the external environment and integrate this information in order to correctly co-ordinate gene expression.

So far, I have studied the regulation of SIGMA FACTOR5 (SIG5) and how it acts upon the *psbD* blue light responsive promoter (BLRP) in the chloroplast in response to environmental and endogenous cues under laboratory conditions. However, to translate this research into an agricultural context it is important to understand the functioning of this mechanism under naturally fluctuating environments.

Therefore, I investigated in the field whether the pathway also operates under naturally fluctuating conditions (e.g. during the day and across seasons) and to identify the main drivers of this pathway in naturally occurring plant populations. For this, we used *Arabidopsis halleri* subsp. *gemmifera*, which is a clonal diploid relative of the annual *Arabidopsis thaliana*. Being an evergreen perennial, it is an excellent model to study plant responses in multiple seasons and has been used successfully to study signalling pathways in the field (or "*in natura"*) (Aikawa et al., 2010;Nagano et al., 2012).

As a model plant, *A. halleri* provides several advantages compared to other Arabidopsis species:

1) It is the closest metal tolerant relative of *A. thaliana*, which generates monocultures in natural habitats polluted with heavy metals. In this case, the *A. halleri* collected for this study grows in rocky soil near a disused copper mine in Japan. This man-created selection allows us to sample only *A. halleri* in the field because it is the only Arabidopsis that can grow there (Kawagoe and Kudoh, 2010).

2) It has high nucleotide sequence identity and good synteny with *A. thaliana*. Sequence analysis of the genes *SIG5*, *psbD* BLRP and *CCA1* from *A. halleri* showed 95%, 89% and 95% identity in nucleotide sequences, respectively, with *A. thaliana* (Hiroshi Kudoh, personal communication).
3) The perennial life cycle allows us to study gene expression all year under natural conditions (Kudoh, 2016).
4) Most of the individuals collected are clones, which homogenises the genetic pool and lowers genetic variation. This is important to reduce noise in expression patterns since the input (environmental cues) is *per se* very noisy.

Light quality through the circadian clock provides spatial and temporal information about the environment. Places under dense vegetation (shaded) are characterized by a low Red to Far Red ratio (R:FR) and if a plant is close to another plant, the reflected light will be perceived by the adjacent plant initiating a process called shade avoidance. Shade avoidance causes developmental alterations that allow plants to overcome possible shading by surrounding vegetation (Franklin, 2008).

Under laboratory conditions, it is known that SIG5-mediated signals communicate information from the circadian oscillator (Noordally et al., 2013), about blue light intensity (Belbin et al., 2016;Onda et al., 2008), R:FR (Belbin et al., 2016) and temperature (Cano-Ramirez et al., under review). To efficiently study the pathway under very complex and usually overlapping environmental stimuli, these previous studies allowed us to narrow the infinite range of cues to monitor to light, light quality and temperature.

We obtained three timecourses of *A. halleri* transcript abundance from a natural population situated in Hyogo, central Honshu (along the Omoide River, Taka-cho); 24th to 26th March 2015, 15th to 17th September 2015, and 13th to 16th of September 2016. In the first two seasons, two sites were selected to sample. In the site designated as sun, the plants received direct sunlight during the day while in the site picked as shade, plants received sunlight filtered by the vegetation most of the time. We sampled leaves from 6 plants randomly located at the sun site every 2 hours between 14:00 to 13:30 hours on the 24th and 25th of March 2015 and on the 15th and 16th of September. For the shade site, we collected leaves from 6 plants situated in different spots every two hours between 14:00 to 14:00 on the 25th and 26th of March and on the 16th and 17th of September. At the same time the sampling occurred, the spectrum and temperature were measured near the plants. To prevent missing points in the data and stress by sampling, four more plants were available at each site to sample. Also, to have a standard for our

qRT-PCR experiments, 10 leaves collected in March 2015 from random plants along the site were used to generate a reference cDNA to compare all the samples from all sampling seasons. In September 2016, local conditions around patches of plants (warming, cooling, light manipulation) were custom built. From all of these experiments I analyzed *SIG5*, psbD BLRP and circadian clock gene expression using multiple linear regression to investigate which environmental cues are driving the pathway under natural conditions.

5.2 RESULTS

5.2.1 *SIG5* promoter has similar activity in *A. thaliana* and *A. halleri*

To evaluate whether the pathway has the same function in *A. thaliana* and *A. halleri* and to determine whether it is appropriate to link the findings *in natura* and laboratory conditions, *SIG5* promoter of *A. halleri* was cloned in a vector containing luciferase to study its activity by bioluminescence assays and compare it with the *SIG5* promoter of *A. thaliana*. Bioluminescence activity screening was tested in T2 and T3 *AhgpSIG5::LUC A. thaliana* plants by vacuum infiltration of luciferin, using a new fast method that I developed (Fig. 5-1).

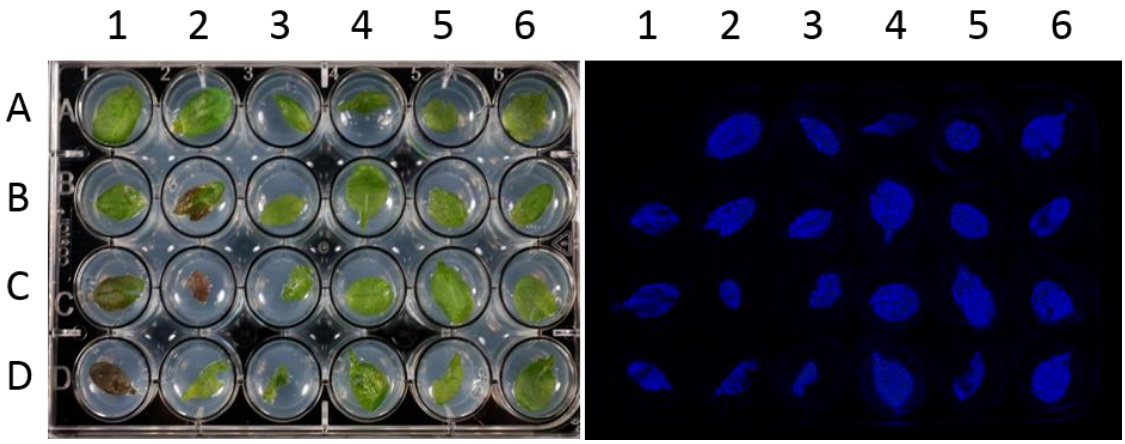


Figure 5-1. Novel luciferase transient expression assay to test positive LUC lines. On the left panel is the plate after dosing with luciferin and on the right is the bioluminescence image. Well 1A is the negative control WT *A. thaliana* and 2A is the positive control *AtSIG5::LUC*. The rest is T3 *AhgpSIG5::LUC A. thaliana* plants.

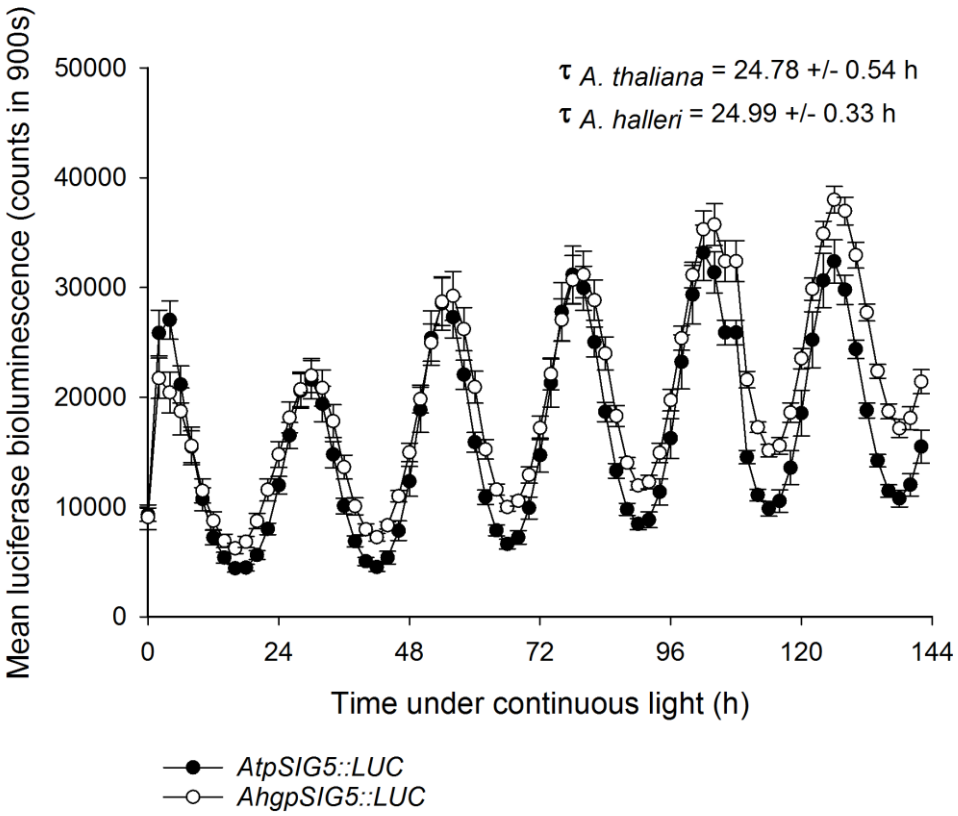


Figure 5-2. SIG5 promoter from *A. thaliana* and *A. halleri* have the same activity and are subjected to similar control by the circadian clock. Time course of luciferase bioluminescence in seedlings expressing *AtpSIG5::LUC* (*A. thaliana*) and *AhgpSIG5::LUC* (*A. halleri*). Photons were counted for 900s every two hours under continuous light for 6 days. Data are means ± SEM (n=6). τ = period in hours.

Afterwards, circadian rhythms of luciferase bioluminescence in *AhgpSIG5::LUC* plants using *AtpSIG5::LUC* as control were monitored. *A. halleri* SIG5 promoter was under circadian regulation, as in *A. thaliana* (Fig. 5-2) with a period length of 24 hours. These results show that *SIG5* in *A. halleri* is under the control of the circadian clock as is *SIG5* from *A. thaliana*. Therefore, we can use this species to study the signalling pathway under natural conditions.

5.2.2 SIG5-mediated signalling to chloroplasts under natural conditions

Sampling was carried out in two seasons, March and September (Fig. 5-3A, 5-3B) in the closest days to equinox (day and night are the closest to 12 h photoperiod). Time of day was explored by sampling leaf tissue every two hours to study whether the pathway is circadian regulated as in laboratory conditions.

Under laboratory conditions, SIG5-mediated signals communicate information concerning light quality (R:FR cue), therefore, the effect of light quality was explored by sampling in two sites on consecutive days, sun and shade, selected based on visual examination of the site (open sky or deep vegetation) and plants (usual phenotype or elongated petioles).

At every time course sampling event, light and temperature were monitored. Light intensity was measured in the field from 200 to 900 nm to obtain a full light spectrum every 5 min. Results indicate that during March, total light intensity was at least four orders of magnitude higher than the total light intensity in September (Fig. 5-3C, 5-3D).

Low temperature is an important driver of the pathway under laboratory conditions (Chapter 4). Therefore, the effect of temperature on the pathway under natural conditions was (initially) investigated by sampling at different seasons. In this case, during March temperatures at the field site ranged from 0 to 17 °C while in September temperatures reached above 20 °C (Fig. 5-3E, 5-3F).

Transcript accumulation was analysed from all samples that could yield good quality RNA, referred to a standard pool of samples collected in March at midday. Clear differences between sun and shade sites only occurred in the period between 26 and 32 hours in March (Fig. 5-4A) and the abundance of *CCA1* transcripts was at least 3 times higher in September at the highest point of accumulation (around 24 h) (Fig. 5-4B).

SIG5 transcript accumulation occurred early at night in the shade site in March (Fig.5-4C) which coincides with colder days and frozen nights recorded at the site (Fig. 5-3A, 5-3E). In September *SIG5* transcripts started to accumulate before dawn, in the absence of light which corresponds to light anticipation under circadian regulation (Fig. 5-4D).

The abundance of *psbD* BLRP transcripts during the day was much lower in the data from March than September, and no clear differences between sites were observed (Fig. 5-4E, 5-4F). Since *SIG5* transcript abundance was not different between sites in September, we did not detect strong differences between sites in *psbD* BLRP induction (Fig. 5-4F).

Overall, the data revealed seasonal differences in the amplitude of gene transcript accumulation in this signalling pathway. Interestingly, our seasonal differences in amplitude of expression coincide with other *in natura* studies. For example,

FLOWERING LOCUS (FLC) homolog in *A. halleri* has the highest transcript abundance around September and the lowest levels of expression in February and March, which coincides with our data for *CCA1* (Kudoh, 2016).

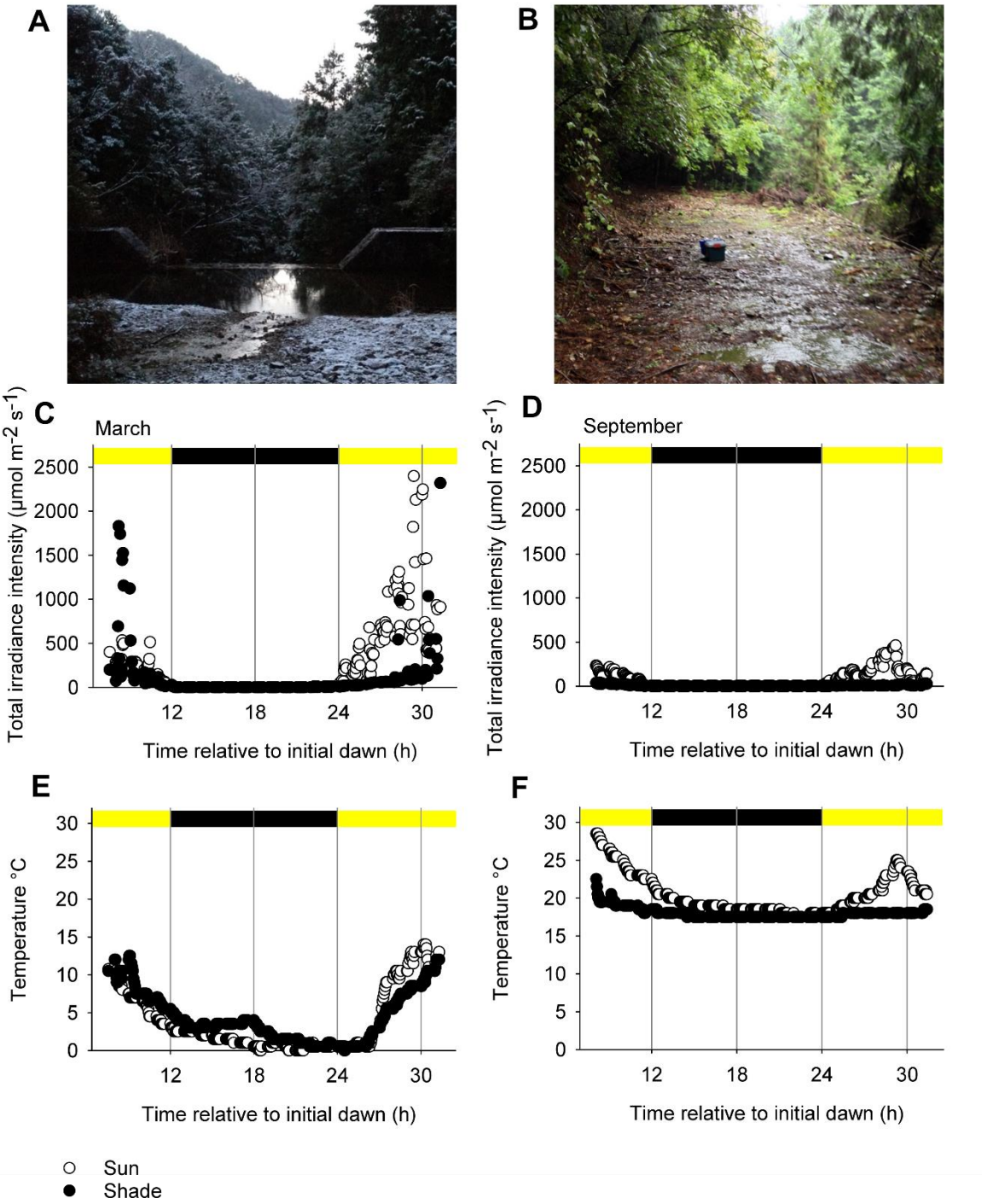


Figure 5-3. Patterns of light and temperature recorded at the natural habitat in March (A) and September (B). Total light intensity from 200-900 nm and temperature measured in March (C, E) and September (D, F). Yellow boxes indicate light period and dark boxes night period. Representative pictures are shown.

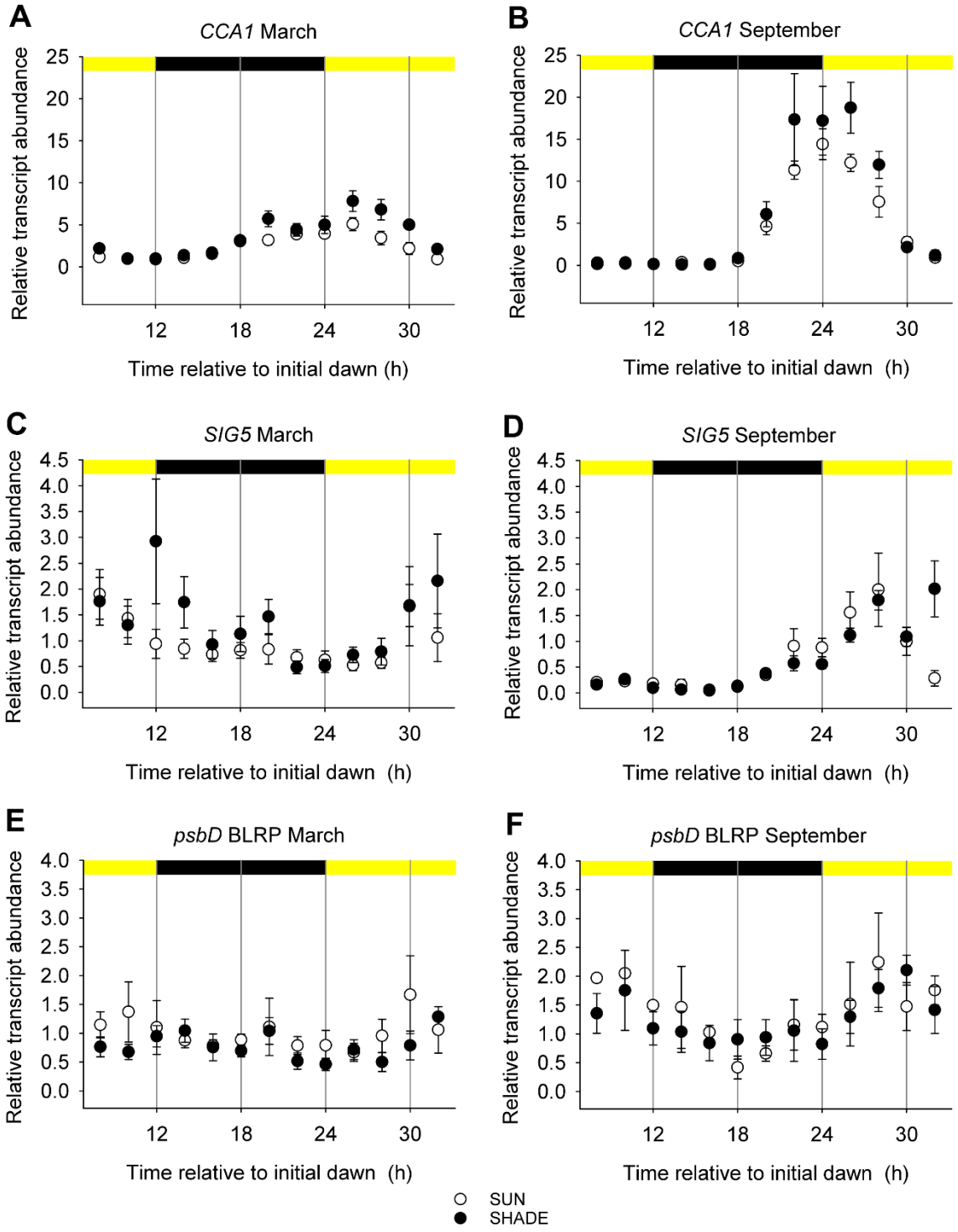


Figure 5-4. SIG5-mediated signalling to chloroplasts in the natural habitat shows distinctive features depending on the light environment and time of year (March and September). Expression analysis of *CCA1* (A, B), *SIG5* (C, D) and *psbD* BLRP (E, F) was determined by qRT-PCR in *A. halleri* leaves collected in sun and shade conditions at the indicated times. Yellow boxes indicate period between sunrise and sunset and dark boxes period between sunset and sunrise. Circles and bars indicate mean ± SE, n=6.

5.2.3 Local manipulation of environmental conditions in the field

These results suggested seasonal differences in amplitude of expression of the genes in the signalling pathway. To separate seasonality from the contribution of each environmental cue, we performed experiments in September 2016 where we manipulated the local environment of groups of plants in the field.

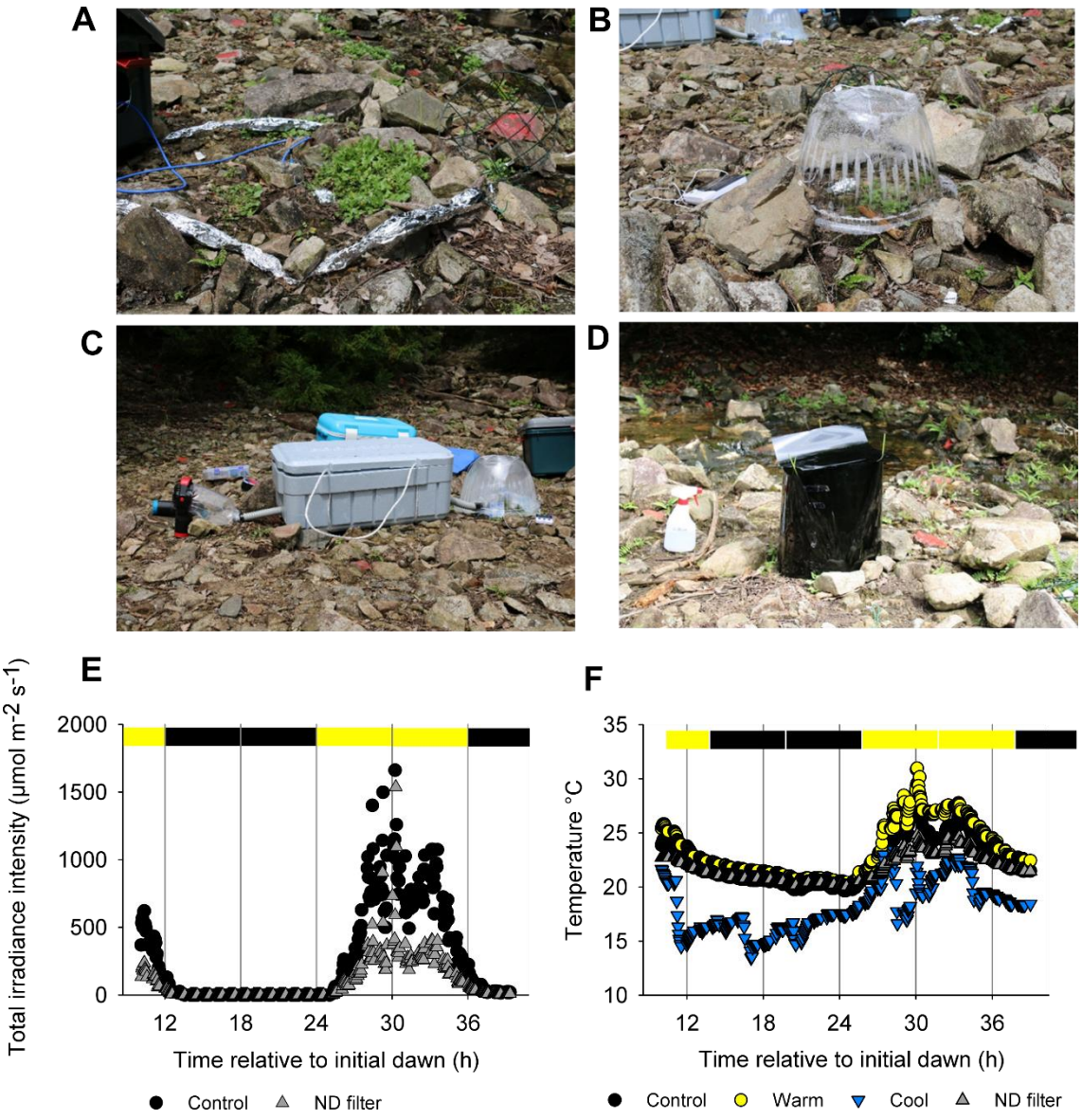


Figure 5-5. Local environment manipulation equipment and patterns of light and temperature recorded at the natural habitat. Total light intensity from 200-900 nm was recorded in control plants (A). Temperature and low light treatments included warm (B), cool (C) and ND filter (D). Total light intensity (E) shown for control and ND filter groups. Temperature was recorded in all sites individually (F). Representative pictures are shown. Yellow boxes indicate period between sunrise and sunset and dark boxes period between sunset and sunrise.

Groups of plants in the area previously known as sun during March and September (2015) were selected. The treatments included a control group (Fig. 5-5A, 5-5E), warm conditions (Fig. 5-5B), cool conditions (Fig. 5-5C) and a neutral density (ND) filter group (Fig. 5-5D, 5-5E).

In general, the light spectrum was reduced to 10% of solar light in the 250 to 700 nm interval (Fig. 5-5E). This is important because some light cues are outside this range. Temperature in the control group was comparable with September from the previous field season (20-30 °C). The warm condition treatment had a modest increase in temperature of around 2°C, while the cool conditions treatment was 5 °C to 10 °C below the control group. The ND filter was slightly cooler than the control group during the light period, with 2 °C difference (Fig. 5-5F).

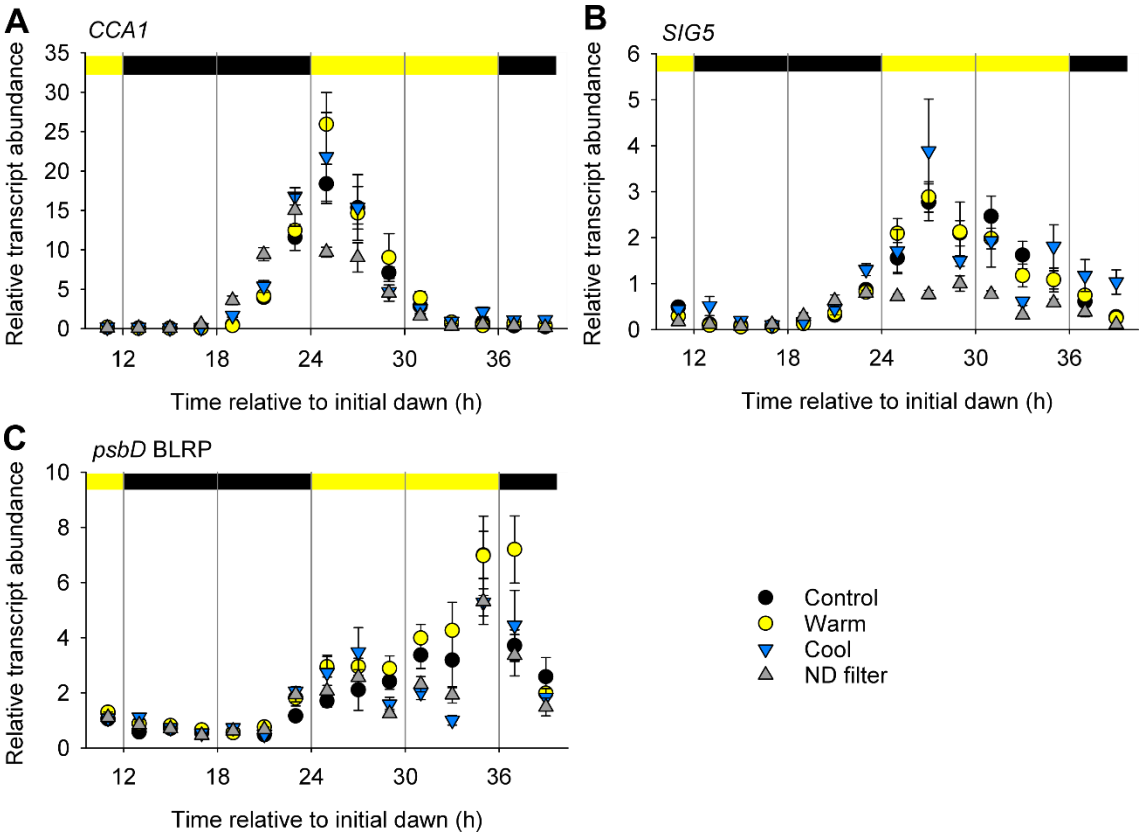


Figure 5-6. Transcript abundance within the *SIG5* pathway was manipulated locally in the field. Expression of (A) *SIG5* (B) *psbD* BLRP (C) *CCA1* determined by qRT-PCR in *A. halleri* leaves collected in control, warm, cool and low light (ND filter) conditions at the indicated times. Yellow boxes indicate period between sunrise and sunset and dark boxes period between sunset and sunrise. Circles and bars indicate mean ± SEM, n=6.

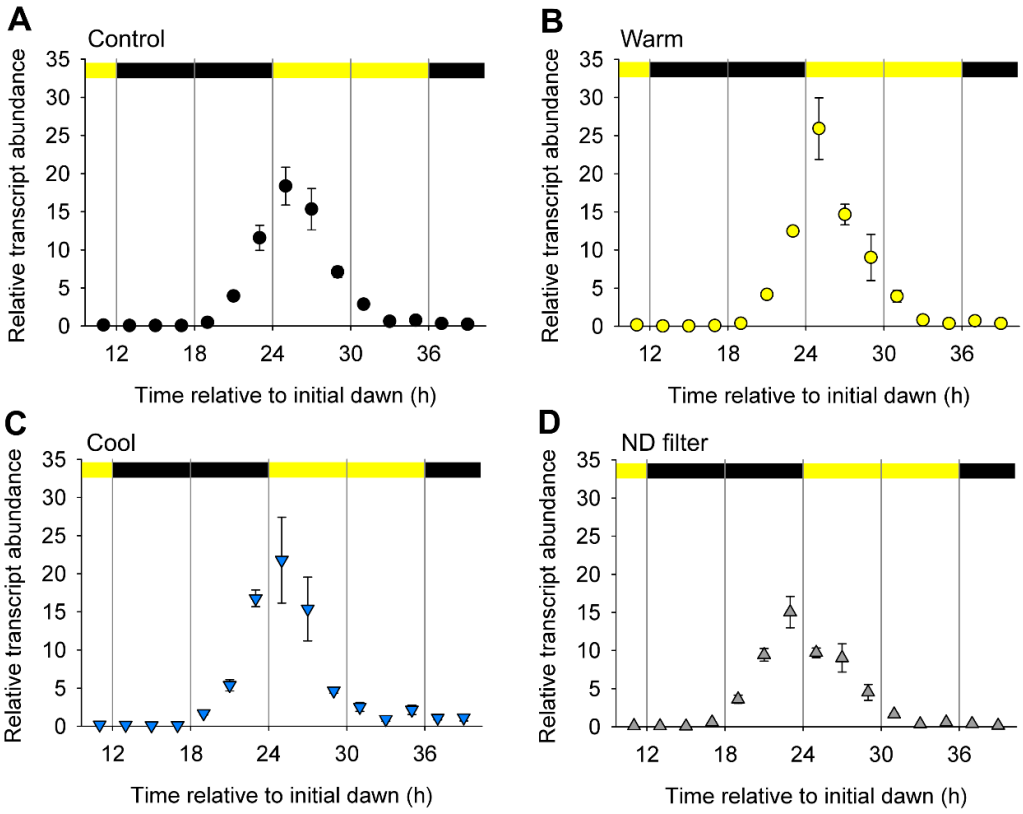


Figure 5-7. *CCA1* peak time depends on light conditions. Transcript accumulation of the morning loop gene *CCA1* from plants locally manipulated in control (A), warm (B), cool (C), and ND filter (D) groups. Yellow boxes indicate period between sunrise and sunset and dark boxes period between sunset and sunrise. Circles and bars indicate mean ± SEM, n=6.

Transcript accumulation was characterised by low abundance during night and induction in anticipation of light signals (Fig. 5-6, before 24h).

Transcript abundance of *CCA1* was comparable to the levels measured during September 2015 in the control group (Fig. 5-7A). Regardless of treatment, transcript abundance increased before dawn and peaked when the plants received light stimuli, except for the plants under the ND filter. In these plants, *CCA1* transcripts peaked before dawn (Fig. 5-7D). This result suggests that the time CCA1 peaks under natural conditions depends on both circadian regulation and light input.

The highest abundance of *CCA1* transcripts occurred in the warm and cool treatments, which might reflect the effects of temperature on the circadian clock.

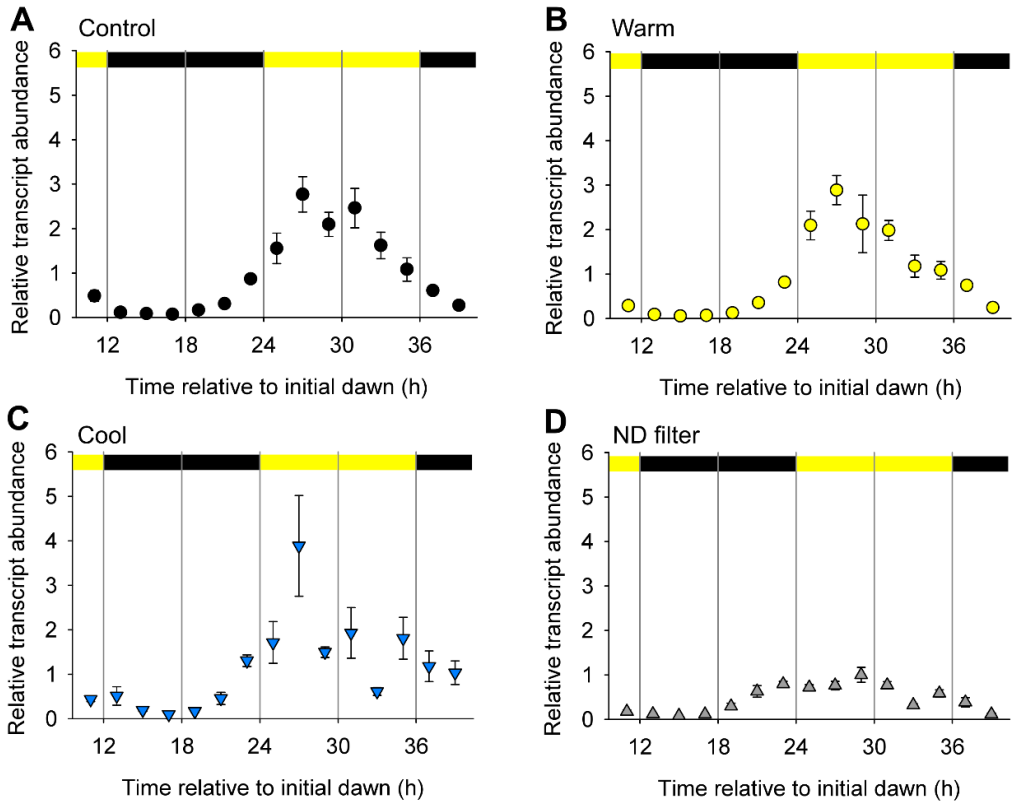


Figure 5-8. *SIG5* transcript accumulation pattern depends on low temperature and light received. Transcript abundance in control (A), warm (B), cool (C), and ND filter (D) treatments. Yellow boxes indicate period between sunrise and sunset and dark boxes period between sunset and sunrise. Circles and bars indicate mean ± SEM, n=6.

SIG5 transcript accumulation in the control group was higher than in previous field seasons (2015), however, in all treatments apart from the ND filter treatment, there was anticipation of dawn, with peak *SIG5* abundance with a plateau in transcript abundance during the light period (Fig. 5-8). The highest abundance of *SIG5* occurred in the cool conditions, in accordance with previous laboratory studies showing that low temperature induces *SIG5* transcript accumulation (Chapter 4).

Chloroplast *psbD* BLRP transcript pattern was strongly influenced by the environmental parameter being manipulated compared to the control treatment (Fig. 5-9).

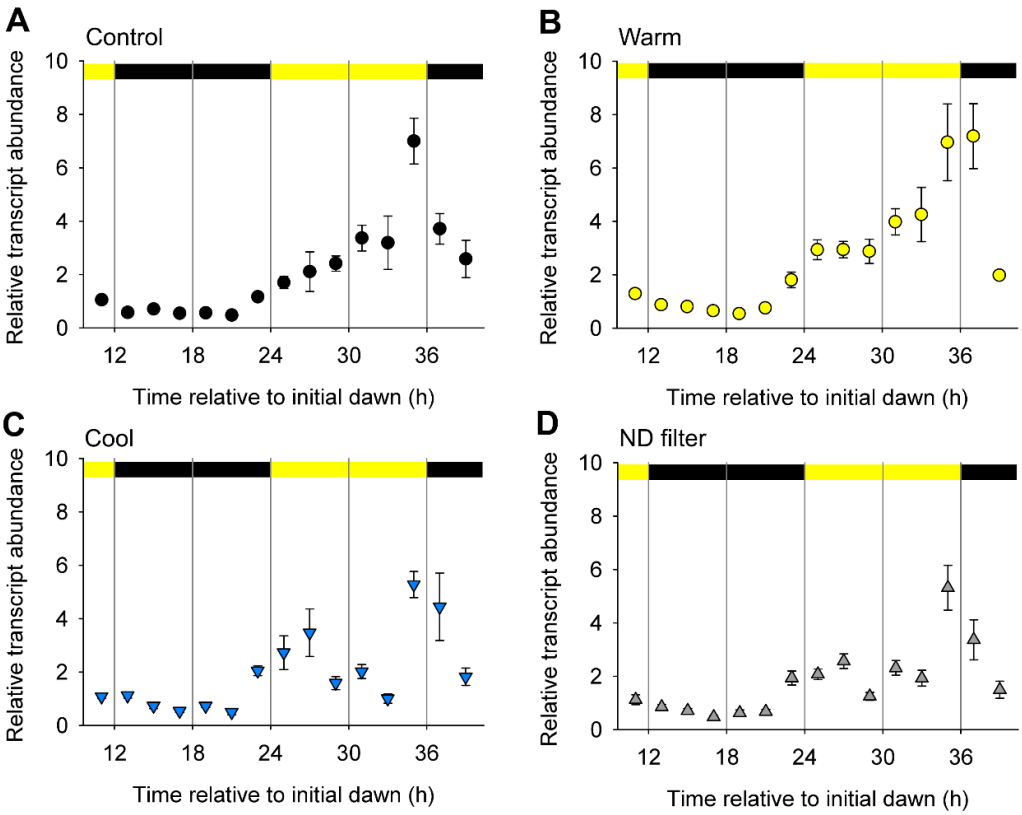


Figure 5-9. *psbD* BLRP transcript accumulation pattern is strongly influenced by environmental manipulations. Transcript abundance in control (A), warm (B), cool (C), and ND filter (D) treatments. Yellow boxes indicate period between sunrise and sunset and dark boxes period between sunset and sunrise. Circles and bars indicate mean ± SEM, n=6.

5.2.4 Signalling pathway dynamic model

Once the data from all field experiments was gathered, the amount of transcript abundance and environmental data was vast. Furthermore, one of the biggest difficulties in analysing data from field experiments is that environmental signals are complex, noisy and the cues affecting expression dynamics between genes are usually simultaneous. Therefore, we used statistical modelling to obtain a comprehensive understanding on which environmental cues are important to communicate to the chloroplast under field conditions and when is the appropriate time. Similar approaches have been used previously in field experiments in rice (Nagano et al., 2012) and *A. halleri* (Aikawa et al., 2010).

After plotting the values of the genes against the following one in the signalling pathway (*CCA1* vs *SIG5*, *SIG5* vs *psbD* BLRP) we found that applying the natural logarithm rendered normal distributions and a linear relationship (Fig. 5-10).

Therefore, I modelled the signalling pathway dynamics between genes using multiple linear regression using the glm function in R (R Core Team, 2014). To build this model, I considered the signalling pathway to be divided in two parts. The first part is the dynamic interaction between *CCA1* and *SIG5*. I thought *SIG5* could be explained by *CCA1* transcript accumulation over time, with the inherent oscillation of *CCA1* providing a time input into the system. I assumed that the information flow was from *CCA1* to *SIG5* since there is no evidence that the information is bidirectional. The second part is the *psbD* BLRP explained by *SIG5*, again we assumed that the information flow was unidirectional from the nucleus to the chloroplast as suggested by earlier work in the laboratory (Noordally et al., 2013).

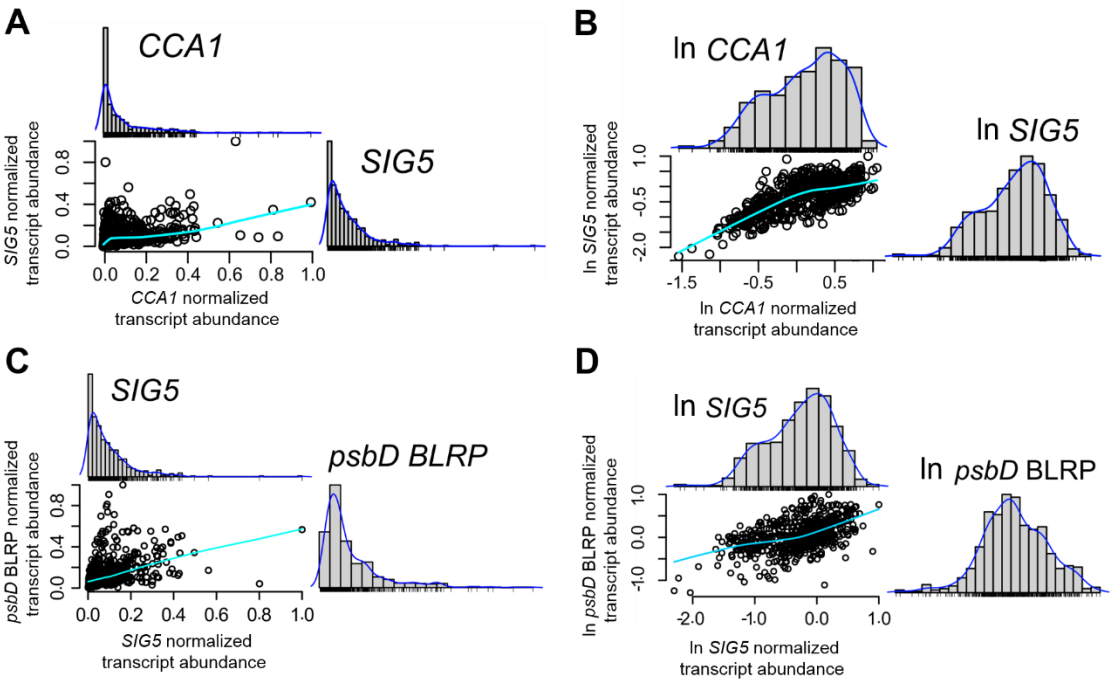


Figure 5-10. Plotting the natural logarithm of the components of the signalling pathway yields a better fit to a linear relationship. Scatterplots of *CCA1* and *SIG5* transcript abundance (A) and ln-transformed values (B). Histogram plots show the distribution of transcript abundance and correspond to the data in the opposite site of the scatterplot. Same plots are shown for *SIG5* and *psbD* BLRP (C, D).

To eliminate the effect of different scales of magnitude in gene transcript abundance, the first step was to normalize then use the natural logarithm (ln) to transform the values:

Equation 1 $$\sigma = \ln(\gamma / \max_{\eta} \gamma)$$

Where σ is the ln-transformed normalized value of the gene transcript abundance (γ) divided by the maximum value across all datasets (η) for each gene.

The simplest model for each system can be described as the following:

Equation 2 $$\sigma_{SIG5} = \sigma_{CCA1} + \beta\varepsilon$$

Equation 3 $$\sigma_{\mathrm{BLRP}} = \sigma_{SIG5} + \beta\varepsilon$$

Where β is the regression coefficient for ε that is the environmental parameter and will be described later.

After normalization, we assembled three datasets combined from multiple sampling seasons, for further analysis:

a) Natural: March and September, sun and shade conditions.
b) Local: The data from the local environmental manipulations (Control, Warm, Cool, and ND filter)
c) Total: Natural and Local datasets combined.

5.2.5 Time delay

Considering that the transcript abundance of three genes was monitored at each timepoint, and there might be a threshold in transcript accumulation for the signal to communicate the information to the next component of the signalling pathway, we would expect that the response of one gene to another within the pathway is not instant. From laboratory data under light dark cycles, *SIG5* peaks 3 h after dawn while *psbD* BLRP peaks 6 hours after dawn (Noordally et al., 2013). Therefore, we wanted to know whether there is also a time delay between the components under natural conditions in the field.

To investigate this, correlation plots were made to determine the most linear relationship between *CCA1* controlling *SIG5*, and *SIG5* controlling *psbD* BLRP. For each system, the values from the second gene were plotted against values from the first gene (predictor) at the same timepoint, the previous timepoint (-2h), two timepoints before (-4h) or three

timepoints before (-6h). Results show that under locally manipulated conditions the best time relationship is a 2h delay between *CCA1-SIG5* and 4h delay between *SIG5-psbD* BLRP. Surprisingly, this relationship changes only under cool conditions for both systems.

In March, *CCA1-SIG5* data had no linear relationship between the components in all delays tested, however, between *SIG5-psbD* BLRP there was a good linearity considering no delay. In September samples had inconsistent optimum delays (Table 1).

Table 5-1. Time delays (h) between the two components of the models specific for each dataset of gene transcription. Values were retrieved from correlation plots on Appendix. Undetermined values are shown as UND.

	Local				Natural			
Gene	Control	Warm	Cool	ND filter	March sun	March shade	September sun	September shade
SIG5	-2	-2	-4	-2	UND	UND	No lag	-2
psbD BLRP	-4	-4	No lag	-4	No lag	No lag	-2	-4

Based on these results, three conditions were tested for building a predictive linear model for the entire dataset. These were no delay, delay through entire dataset or best correlation delay only applied to relevant conditions.

To assess how well our models can describe and predict the experimental data, we assessed them by plotting the experimental transcript abundance and predicted transcript abundance derived from the linear model. The coefficient of determination (R^2) and the Hosmer Lemeshow goodness of fit test were used to quantitively evaluate the models. The higher the R^2 indicates a better linear relationship between the experimental and predicted transcript abundance while the Hosmer Lemeshow goodness of fit test evaluates if there is a statistical difference between the experimental and predicted datasets ($p<0.05$) so the higher p-value, the better the fit.

Results show that R^2 increased when a delay was incorporated into the *CCA1- SIG5* model, however it had the same value when a general -2 h lag between genes was applied or the optimum lag for each condition was evaluated. Therefore, based on the highest p-value, the -2 h lag dataset was selected for further analysis (Fig. 5-11).

For *SIG5-psbD* BLRP, the highest R^2 value arose when a customized time delay was used for each experimental condition, therefore this dataset was selected for use in subsequent analysis (Fig. 5-12).

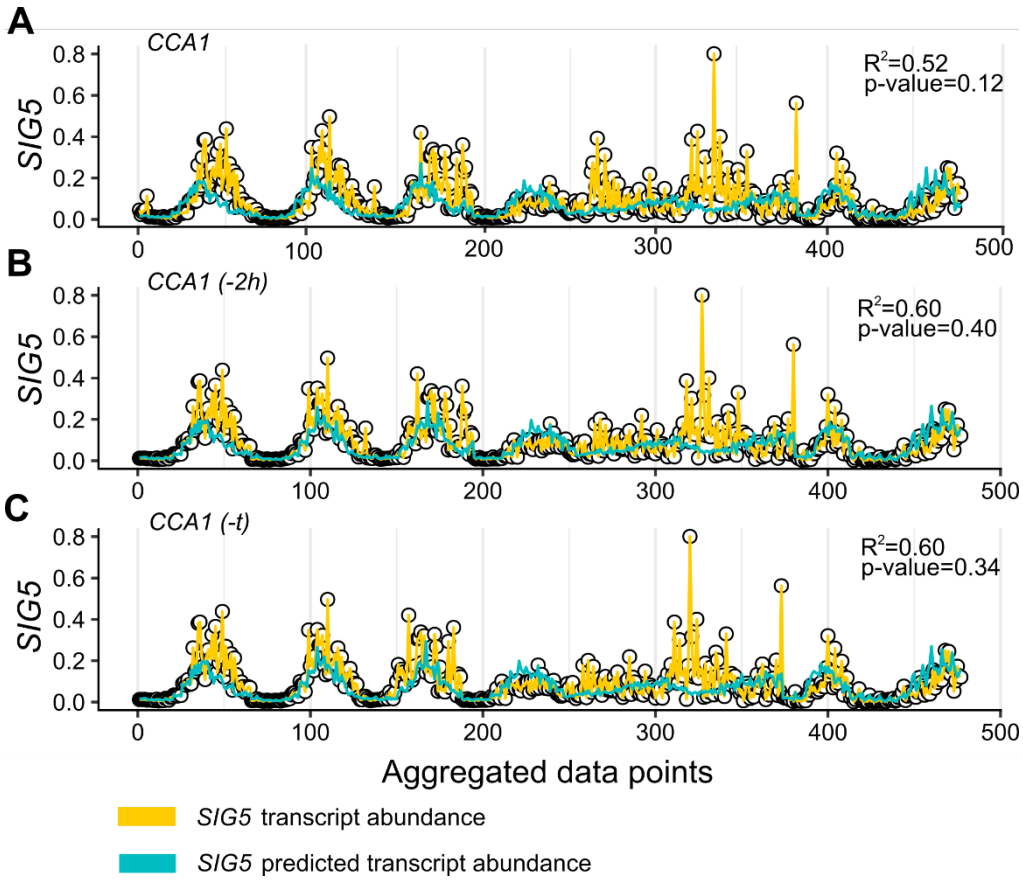


Figure 5-11. Evaluation of different delays between *CCA1* and *SIG5* in the total transcript abundance. Analysis was performed on datasets where *CCA1* and *SIG5* had no delay (A), a general -2h delay (B) or customized delay according to Table 1 (C).

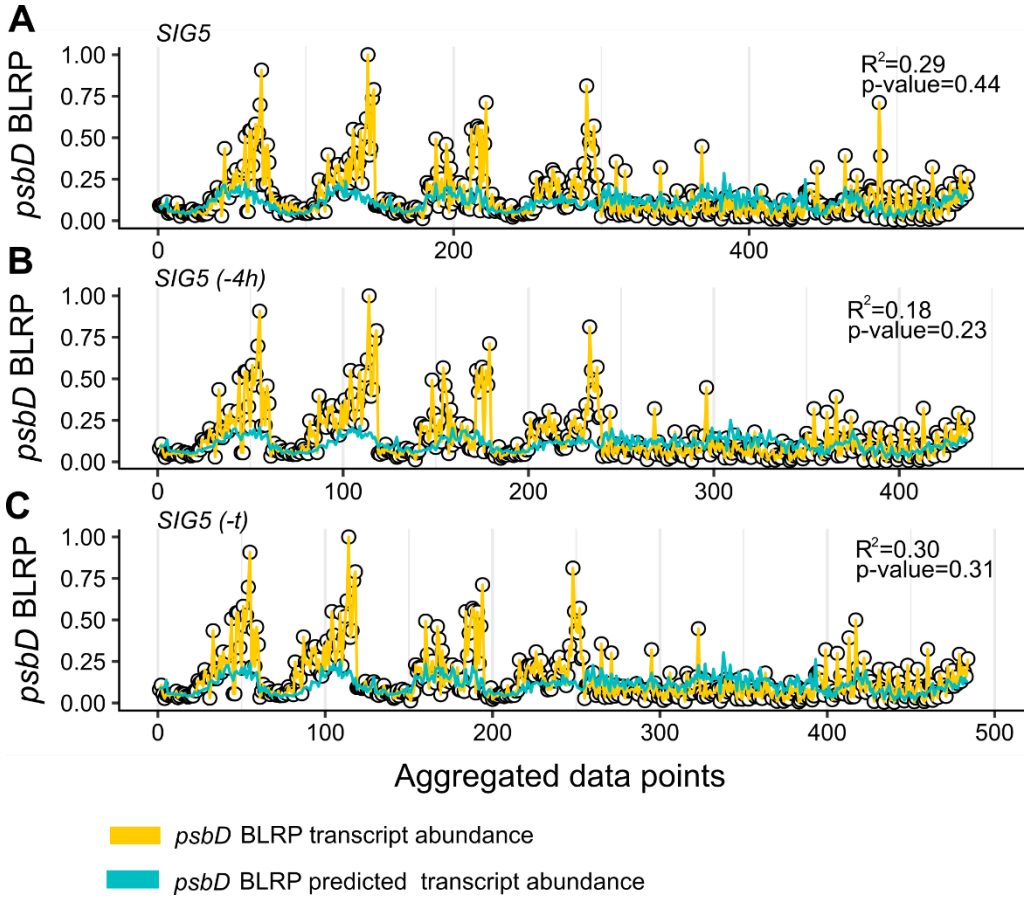


Figure 5-12. Evaluation of different delays between *SIG5* and *psbD* BLRP in the total transcript abundance. Analysis was performed on datasets where *SIG5* and *psbD* BLRP had no delay (A), a -4h delay applied to all dataset (B) or customized delay according to Table 1 (C).

5.2.6 Environmental data

To investigate the main environmental drivers of the accumulation of these transcripts, we incorporated single environmental parameters into the models in the three datasets (local, natural and total) with their respective delays between genes.

We considered there are two main components of environmental stimuli, light and temperature. As light was measured as the complete spectrum between 200-900 nm, we could derive several parameters:

a) Total light intensity calculated as the integral of light intensity between 200-900 nm.
b) Photosynthetically active radiation (PAR), calculated as the integral of light intensity in µmol/m2/s over the 400-700 nm range.
c) Red light in the 660-670 nm wavelength.

d) Far red-light (FR) in the 725 -735 nm wavelength.
e) R:FR, calculated as the ratio of photon irradiance between red and far red light.
f) Blue light intensity in the 390 and 440 nm range.
g) UV-B light intensity between 280 and 315 nm.

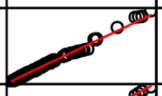
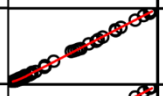
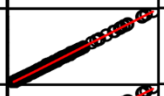
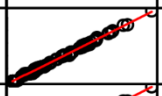
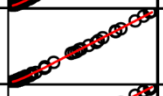
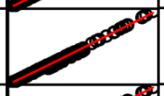
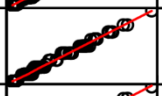
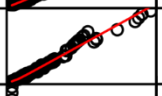
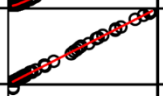
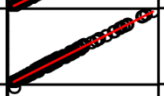
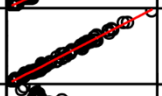
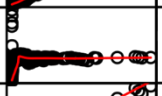
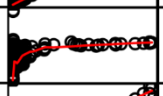
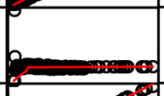
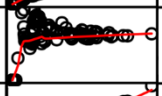
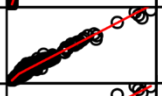
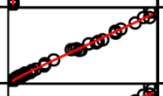
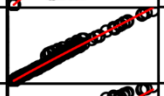
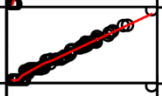
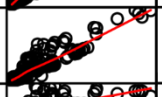
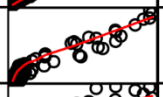
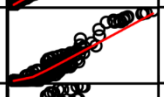
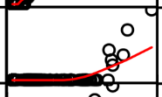
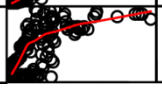
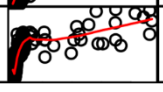
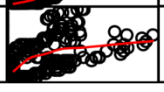
The resulting eight environmental parameters values (seven from light, plus temperature) are in very different scales, therefore, to consider all of them in the models we normalized the data with the following function:

Equation 4 $$\varepsilon = {\delta}\Big/{\max_{\eta} \delta}$$

Where ε is the normalized value of the environmental parameter value from the moving average of the previous two hours before the sampling point (δ) divided by the maximum value across all datasets (η) for each parameter.

One important consideration when assessing contributions of environmental cues was that only non-correlated cues can be added sequentially to the model. In our study, the information that can be extracted from the light spectrum contains other information, for example total light intensity includes PAR and the latter contains all visible wavelengths (Table 5-2). As seen in the correlation plots between the parameters, some contain the same information; PAR, red, far red, and blue light are the same as total light intensity (correlation equals ~1, Table 5-2). To avoid over fitting due to additive effects, we decided to choose a smaller portion of the total light spectrum as light input, and total light intensity and PAR were not considered as major determinants in the following tests due to their broad coverage.

Table 5-2. Correlation table between the eight environmental parameters selected for analysis.

Environmental parameter	Total light March Sun	Corr.	Total light March Shade	Corr.	Total light September Sun	Corr.	Total light September Shade	Corr.
PAR 400-700 nm		1.00		1.00		1.00		1.00
Red 660-670 nm		0.99		1.00		1.00		1.00
Far Red 725 -735 nm		0.99		1.00		1.00		1.00
R:FR		0.43		0.43		0.60		0.69
Blue 390 - 440 nm		0.98		1.00		1.00		0.99
UV-B 280 - 315 nm		0.94		0.94		0.95		0.47
Temperature °C		0.77		0.68		0.61		0.78

5.2.7 Model evaluation adding the environmental input terms

Previously, model outcomes were assessed by R^2 and the goodness of fit statistical analysis because the datasets had different lengths (a -2h lag shortens the dataset). However, now that all datasets have the same number of points, models were evaluated using the Akaike information criterion (AIC). This is an estimator of model quality where a lower AIC means a better fit while a high value indicates low quality model. This is widely used as a model selector and to reduce parameters (Wagenmakers and Farrell, 2004).

Results from the analysis (AIC and R^2) of three datasets (local, natural and total) were plotted and major environmental determinants were selected based on lower AIC values and higher R^2.The predictive performance of the model for *CCA1-SIG5* was the best when R:FR, temperature or UV-B data was considered in the local, natural and total datasets respectively (Fig. 5-13A). In the case of *SIG5-psbD* BLRP, the model was improved by inclusion of temperature in the local, natural and total datasets.

However, R^2 values which quantitatively explain how much of the experimental data can be attributed to the terms in the equation, are below 0.5 (50%) for *SIG5* and only 0.2 (20%) for *psbD* BLRP in the natural dataset (Fig. 5-13). Furthermore, the addition of the environmental parameters improves the AIC values, however R^2 remains the same.

To understand this in more depth, we plotted the total dataset and the predicted values and found that the area that corresponds to the March sun and shade datasets (between 300 and 400 marks in x, Fig. 5-14A), the model performance is very poor as there is little overlap between the experimental and predicted data when examined closely (from 0 to 100 in x, Fig. 5-14B).

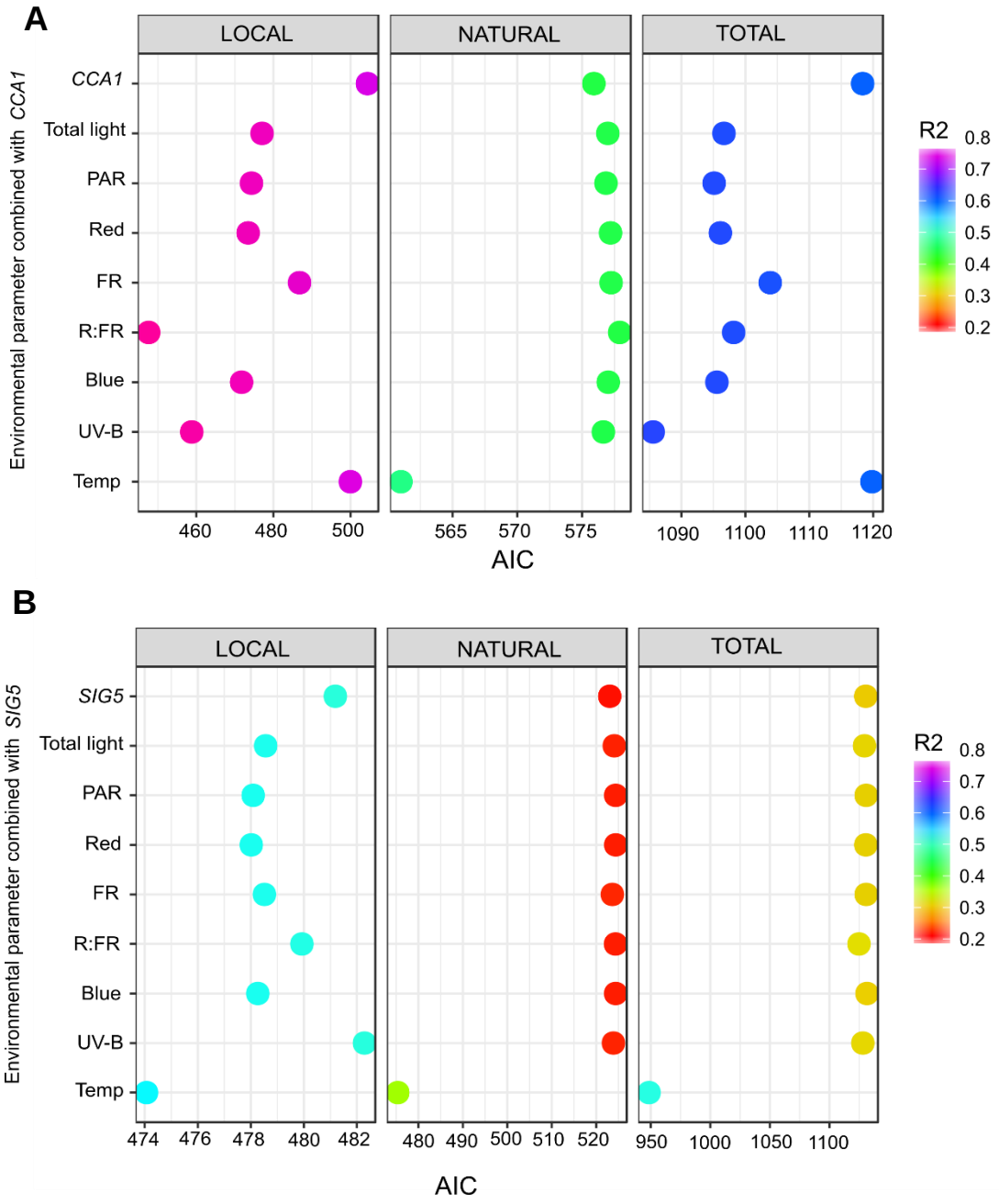


Figure 5-13. Environmental input terms improved model quality. Environmental parameters lowered the AIC value when added to the *CCA1-SIG5* (A) and *SIG5-psbD* BLRP (B) models but R^2 increased only in the latter in natural and total datasets.

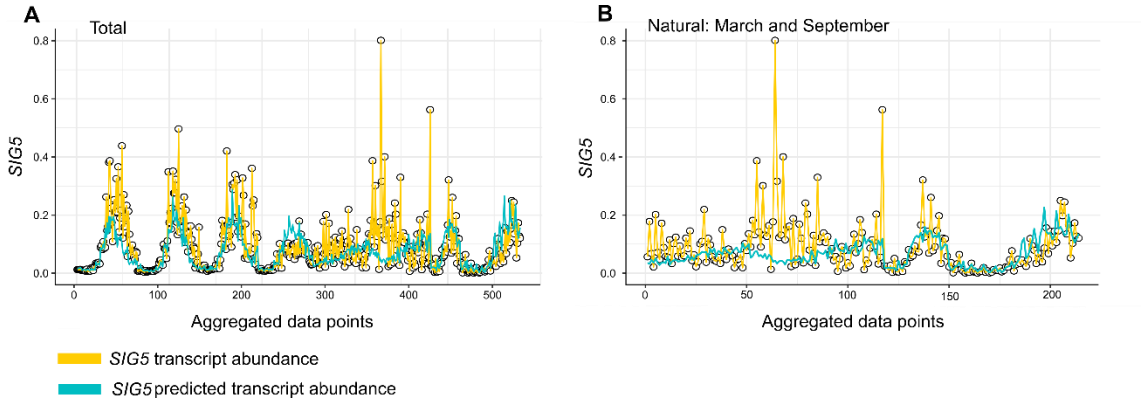


Figure 5-14. Model failed to predict natural dataset. Scatterplot of the experimental transcript abundance and predicted values in total (A) and natural (B) datasets.

Examining the datasets from March and compared to the September datasets (Fig. 5-3A, 5-3B), *CCA1* peak magnitude is 30% less than *CCA1* peak in transcript abundance in September, which suggest the *CCA1* oscillations during March are damped. Previous studies have shown that the sensitivity of this pathway to environmental stimuli changes during the day and this is controlled by the circadian oscillator (Noordally et al., 2013).

We reasoned that a low amplitude clock during colder months might lead to no input from the central oscillator to *SIG5*. Due to this, we wanted to test whether removing oscillatory inputs for the March period improved the predictive quality of the model. If so, this would indicate that the sensitivity to environmental stimuli depending on time of day is ignored or nulled in March.

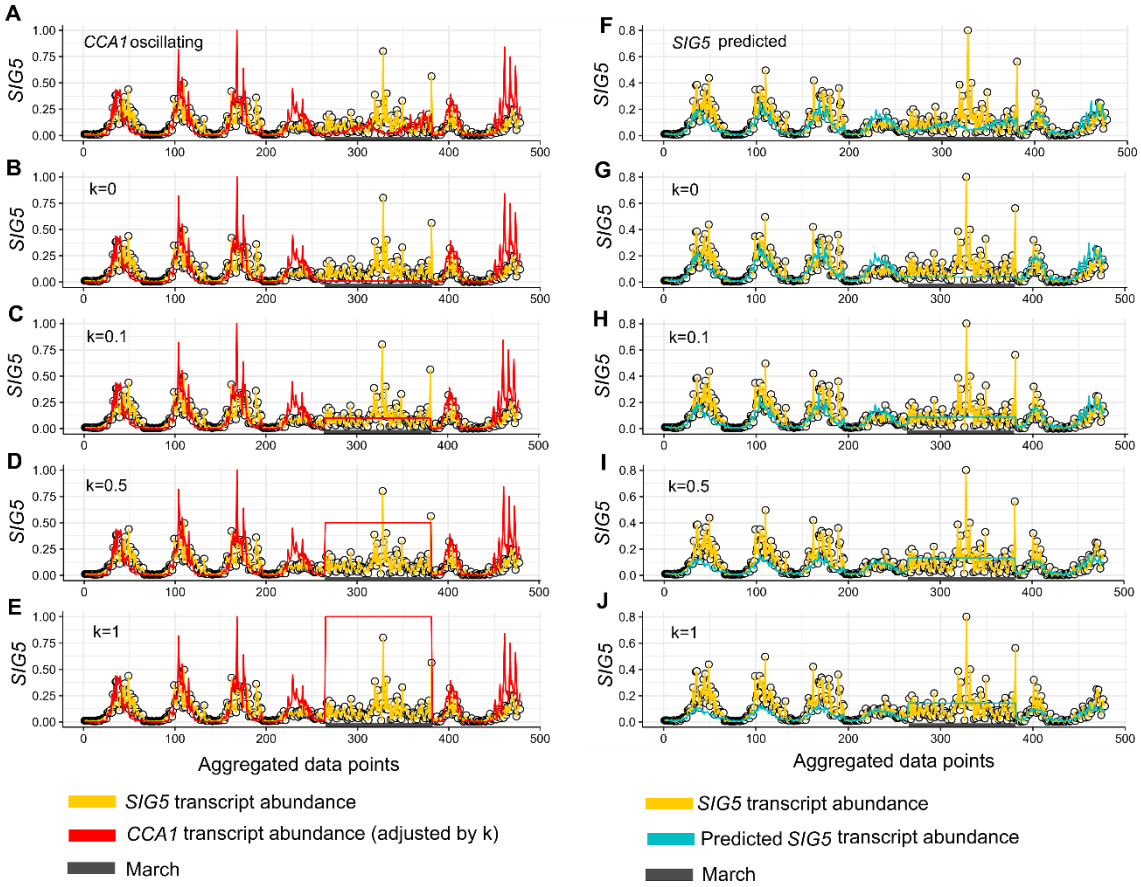


Figure 5-15. Evaluation of models where *SIG5* is uncoupled from *CCA1* oscillations in March. We compared the original *CCA1* transcript abundance and predicted output (A, F), with models incorporating continuous values (between 0 and 1) for CCA1, 0 (B, G), 0.1 (C, H), 0.5 (D, I) and 1 (E, J). k is constant value set for *CCA1*.

We compared the predicted values from the original experimental transcript abundance in the total dataset of *SIG5* to the outcome from the models considering four continuous values in the scale from 0 to 1 for *CCA1*. Replacing the oscillating *CCA1* values from the data in March (sun and shade) with constant values of 0, 0.1, 0.5 or 1 yielded rectangular windows of predicted *SIG5* values with different heights (Fig. 5-15). Further analysis combining the different values given to *CCA1* with the environmental parameters in the model was carried to evaluate AIC and R^2.

Compared to the original oscillating transcript abundance data for *CCA1*, setting constant values of 0.1 for *CCA1* in March renders the best quality model as determined by AIC, furthermore, adding the environmental parameter term increased the R^2 value (Fig. 5-16).

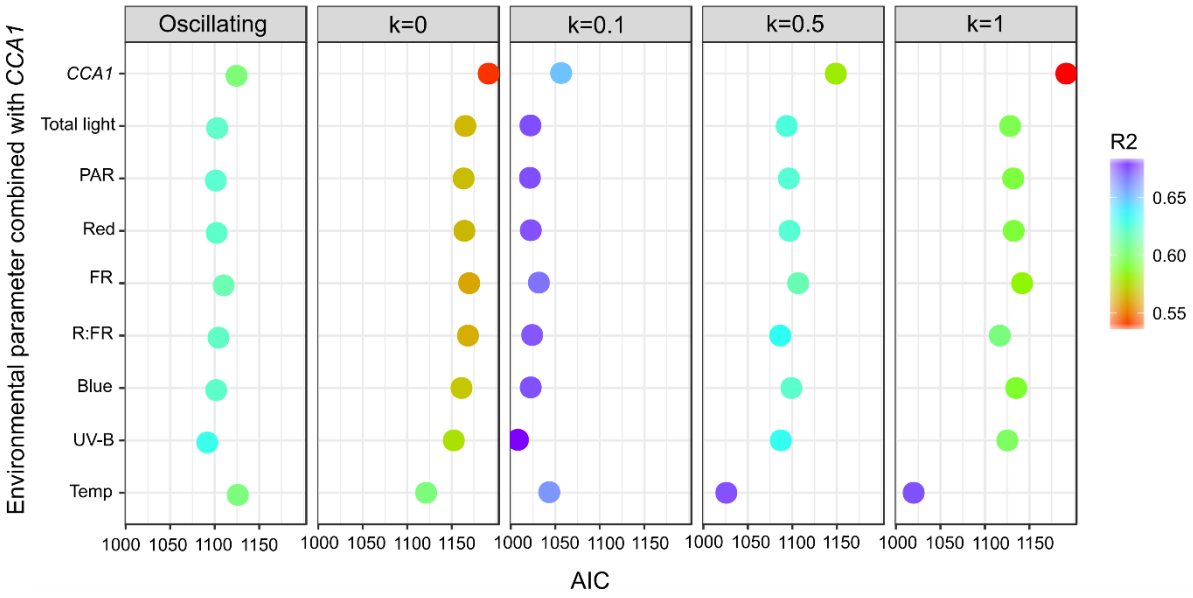


Figure 5-16. *SIG5* transcript abundance is better explained when *CCA1* input during March is not oscillating and has low amplitude. AIC and R^2 values extracted from the analysis from Fig. 5-15 of five different scenarios of *CCA1* fed to the model.

5.2.8 Major determinants of the signalling pathway

To this point, we have optimised the model by incorporating a delay between the two genetic components of each model and by removing the oscillatory effect of *CCA1* in March data.

To investigate which environmental stimuli control the signalling pathway, both models *CCA1- SIG5* and *SIG5-psbD* BLRP were analysed by including the previous optimisation steps and analysing the local, natural and total datasets. The environmental parameter that best explained the dynamics between *CCA1- SIG5* for the locally manipulated plants was R:FR, while for the total dataset was UV-B. In the dynamic between *SIG5-psbD* BLRP, temperature was the main driver in all datasets (Fig. 5-17).

Further analysis was carried out by plotting the experimental and predicted gene transcript abundance with the top four non-correlated environmental parameters (R:FR, blue, UV-B, and temperature) that improved the quality of the models. The model without environmental term for *CCA1-SIG5* indicates that 75% of the experimental values can be explained by *CCA1* transcript abundance (Fig. 5-18F). When the environmental term was added to the model, R:FR had the highest reduction in AIC (56 units) and improved fit by 4% (Fig. 5-18G). R:FR was followed by UV-B (Fig. 5-18H), blue (Fig. 5-18I) and red light (Fig. 5-18J) as environmental stimuli term.

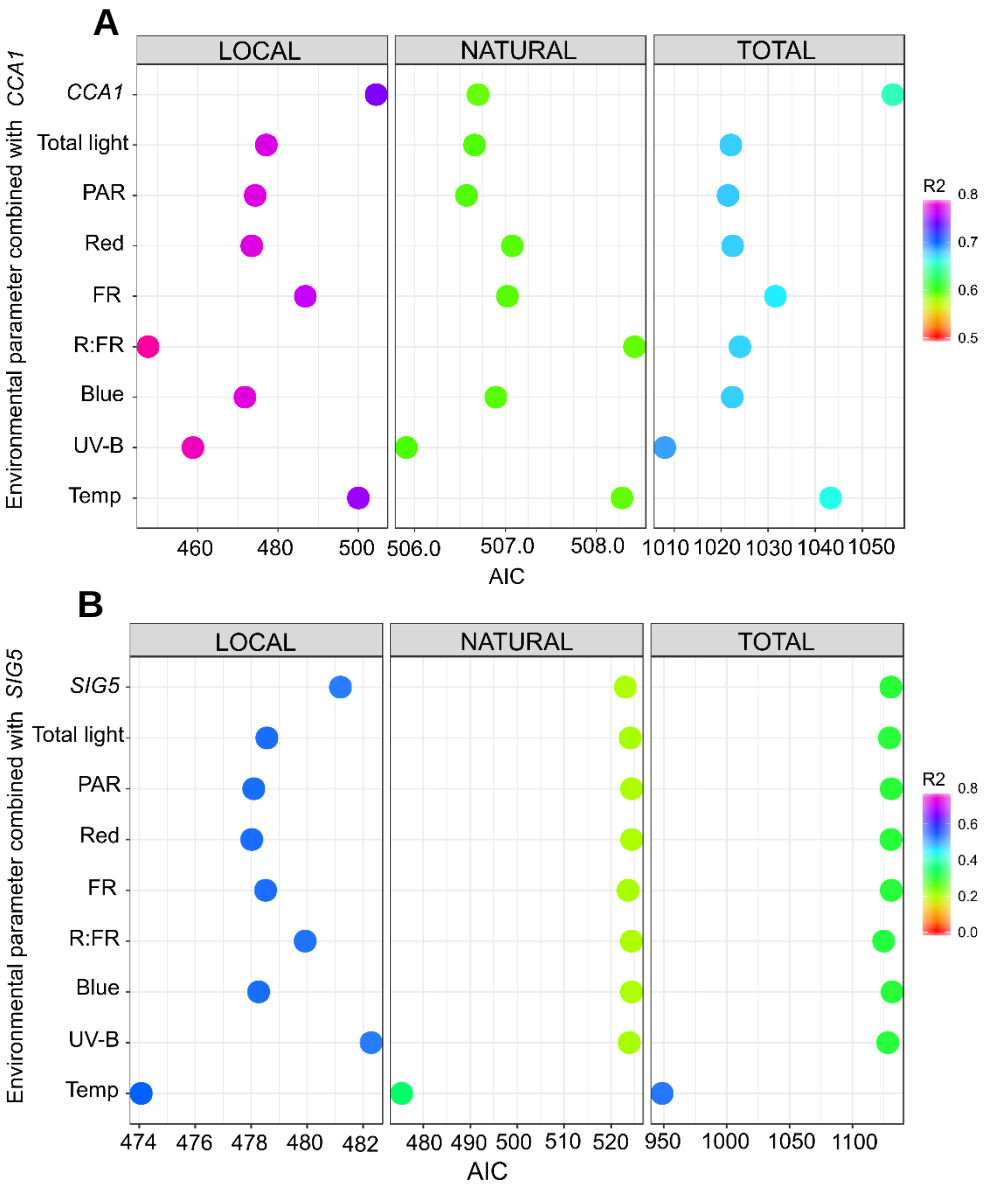


Figure 5-17. Major environmental drivers of the signalling pathway include R:FR, UV-B and Temperature. For the locally manipulated plants, the main driver is R:FR while total datasets indicate it is UV-B for *CCA1-SIG5* (A). Temperature was the main environmental cue for *SIG5-psbD* BLRP (B).

Since we used the constantly set values for *CCA1* in March in this analysis, a deeper analysis revealed that in the natural dataset the environmental parameter that explained the gene expression dynamics was UV-B. However, the addition of UV-B parameter in the model results in a lower AIC value by less than two units, which is not considered as an improvement of the model (Burnham and Anderson, 2004). Therefore, in the natural dataset the transcript abundance of *CCA1* by itself explains 59% of *SIG5* transcripts and the environmental term is unable to explain the remaining 41% (Fig. 5-19).

For the complete dataset (total) inclusion of UV-B, blue, R:FR and temperature individually improved the model as identified by a reduction in the AIC values (in a range of 48 to 13 units, Fig. 5-20), however UV-B addition caused the greatest reduction in AIC (Fig. 5-20G).

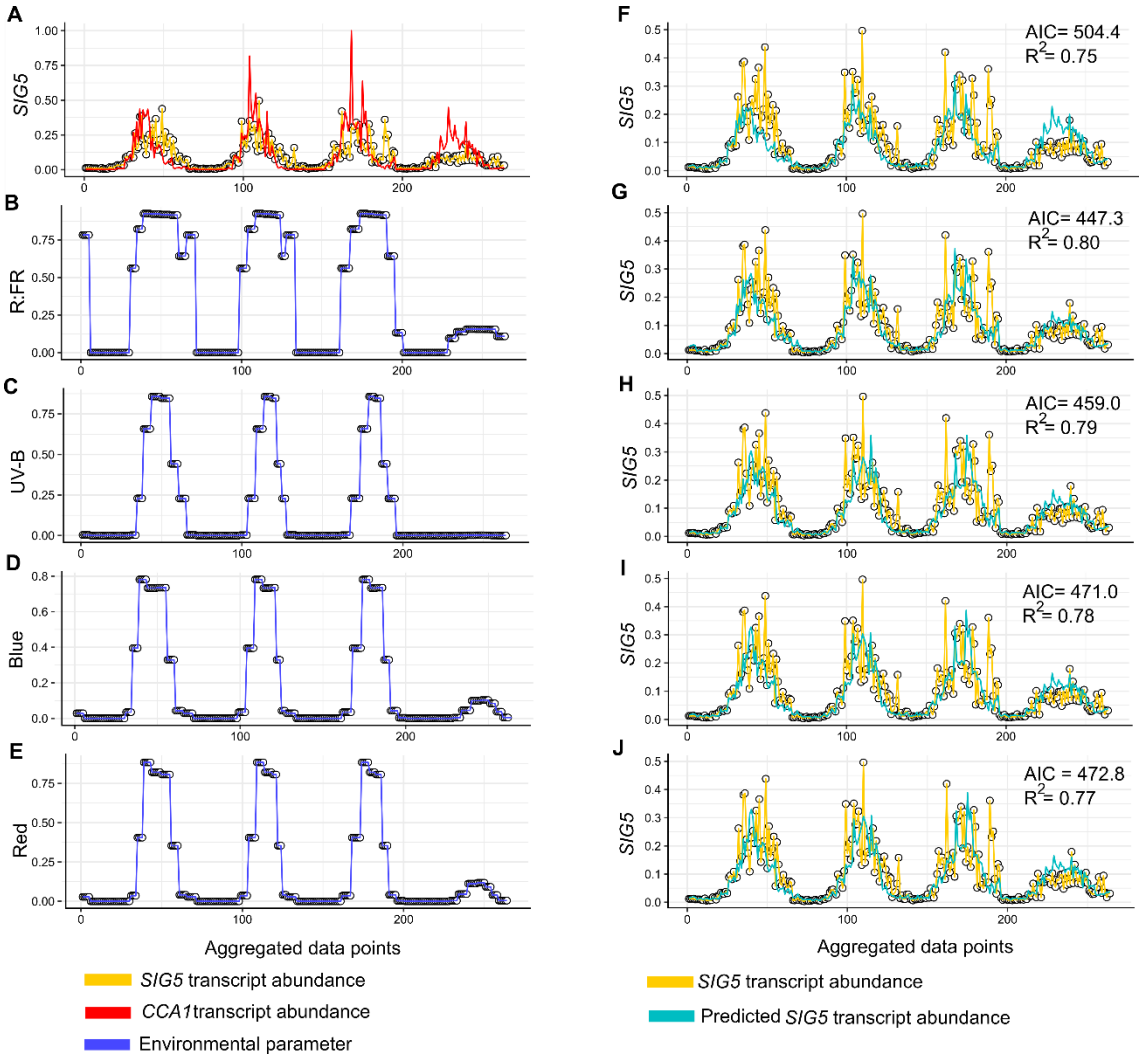


Figure 5-18. Comparison of *SIG5* experimental and predicted transcript abundance in the local dataset with the environmental term added. The top four environmental cues that individually improve the *CCA1-SIG5* model (A, F) are R:FR (B, G), UV-B (C, H), blue (D, I) and red light (E, J).

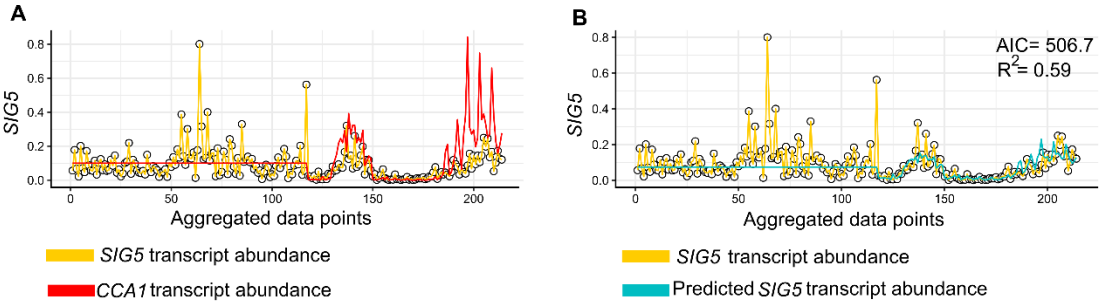


Figure 5-19. In the natural dataset no environmental term improved the *CCA1-SIG5* model (A, B).

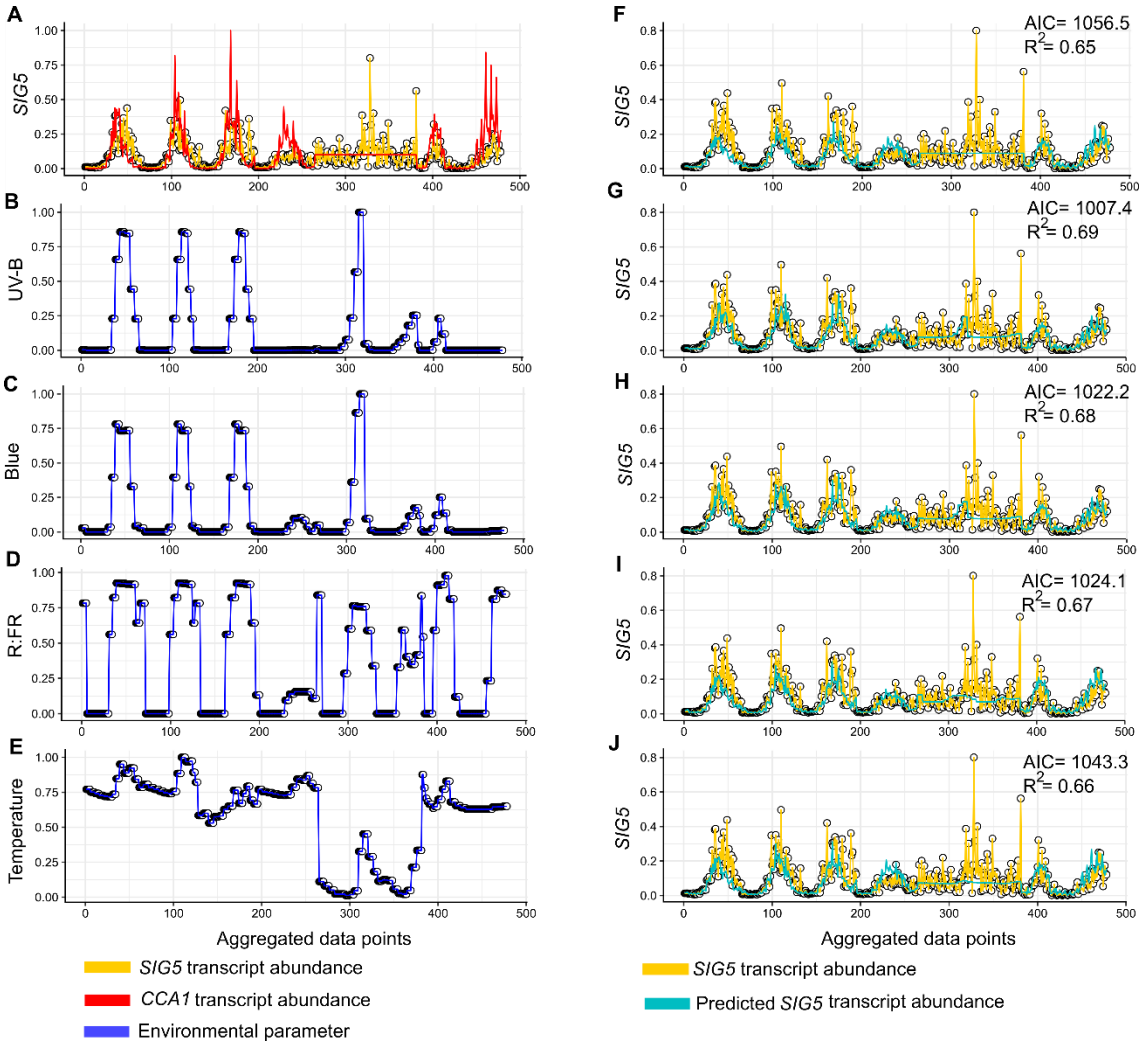


Figure 5-20. Comparison of *SIG5* experimental and predicted transcript abundance with the environmental term added (total dataset). The top four environmental cues that individually improve the *CCA1-SIG5* model (A, F) are UV-B (B, G), blue (C, H), R:FR (D, I) and temperature (E, J).

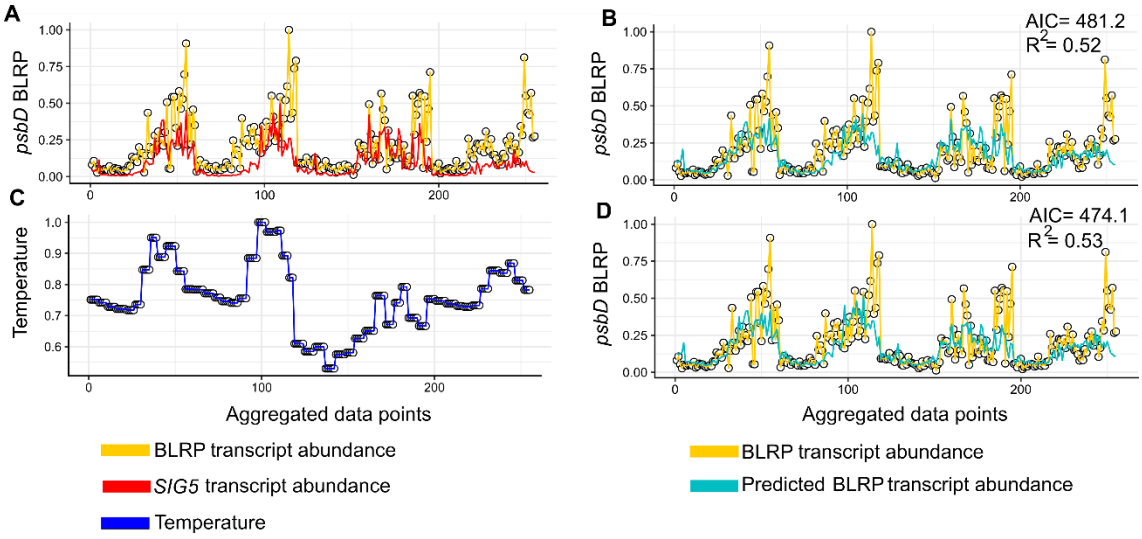


Figure 5-21. Inclusion of temperature improves the prediction of *psbD* BLRP transcript abundance in the local dataset for the *SIG5-psbD* BLRP model.

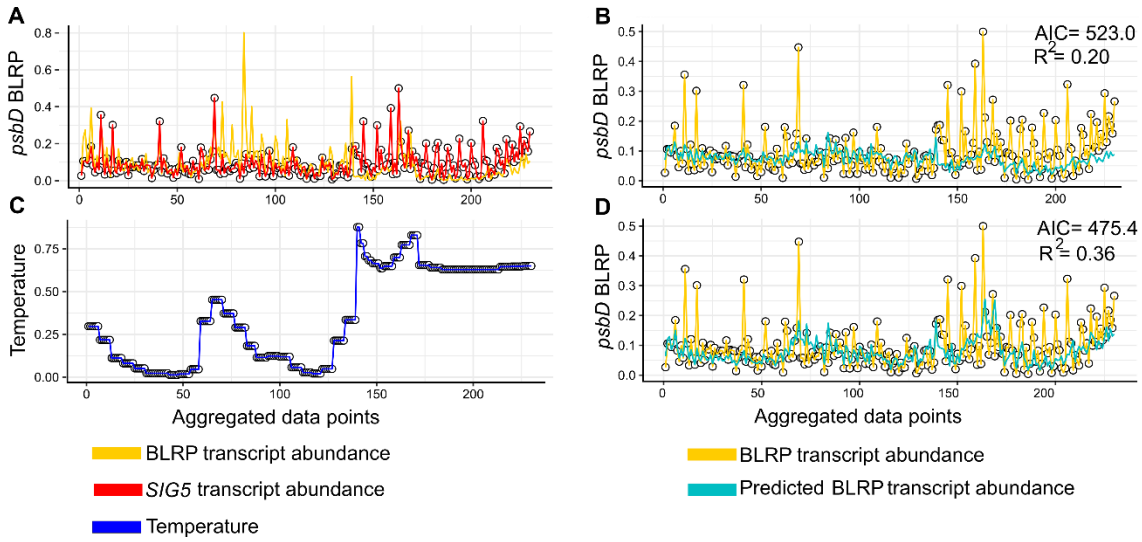


Figure 5-22. Inclusion of temperature improves the prediction of *psbD* BLRP transcript abundance in the natural dataset for the *SIG5-psbD* BLRP model, but it fails to predict 65% of *psbD* BLRP transcript abundance.

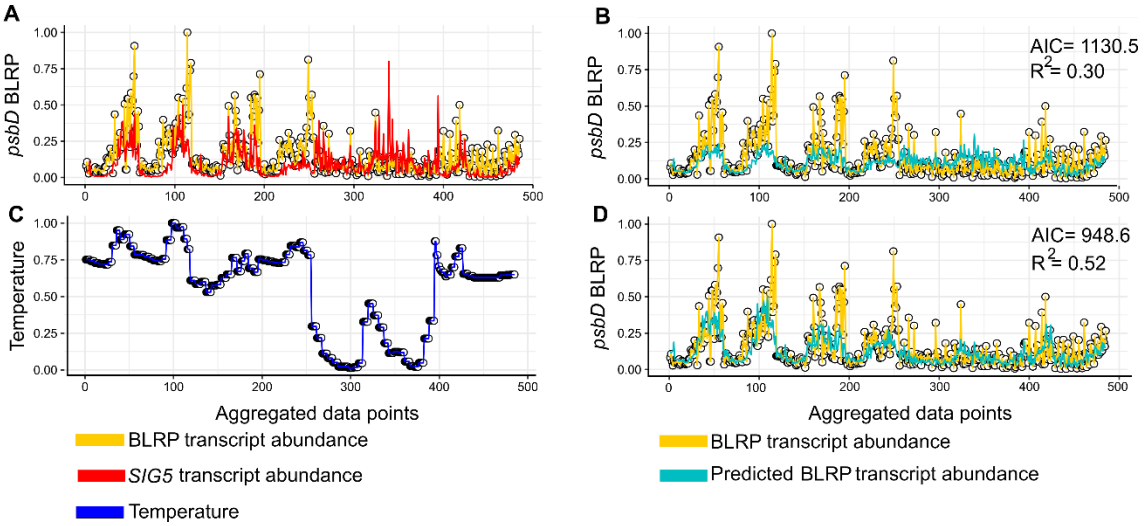


Figure 5-23. Inclusion of temperature improves the prediction of *psbD* BLRP transcript abundance in the total dataset for the *SIG5-psbD* BLRP model.

The *SIG5-psbD* BLRP model indicates that up to 52% of *psbD* BLRP transcript abundance can be explained by *SIG5* transcript abundance and temperature for the local environment manipulation (Fig. 5-21) and total (Fig. 5-23) datasets. Unfortunately, only 36% of *psbD* BLRP transcript abundance can be attributed to *SIG5* and temperature in the natural dataset (Fig. 5-22), so the model failed to explain 64% of the chloroplast gene transcript abundance.

5.2.9 The role of R:FR in the local environmental manipulation experiments

R:FR was the principal environmental cue that explains part of the *CCA1-SIG5* signal transduction from the local dataset analysis (Fig. 5-17A). Nevertheless, after analysing the pattern for each cue we noticed that the environmental stimuli have an approximately square-wave pattern, unlike any other light cue (Fig. 5-18B). Furthermore, due to the way R:FR was calculated, at night red and FR light intensities were zero, but zero divided by zero is undefined so we set the night R:FR values to zero.

Therefore, I wanted to further investigate whether the contribution of R:FR is due to the actual ratio, or day and night information. To do this, all dark period information was removed from the three datasets, with darkness defined as PAR intensity zero, and the analysis was repeated.

For the *CCA1-SIG5* relationship, R:FR and UV-B explained most of the gene expression dynamics in locally manipulated plants after removal of the dark period data. Interestingly, the natural dataset shifted from no environmental input to R:FR or UV-B as

the main driver of the pathway in the absence of night information. In total, the major environmental explanatory term for the whole dataset was R:FR, however, UV-B was the second best environmental cue in decreasing the AIC value of the model (Fig. 5-24A). Further data analysis is needed in order to clarify these results.

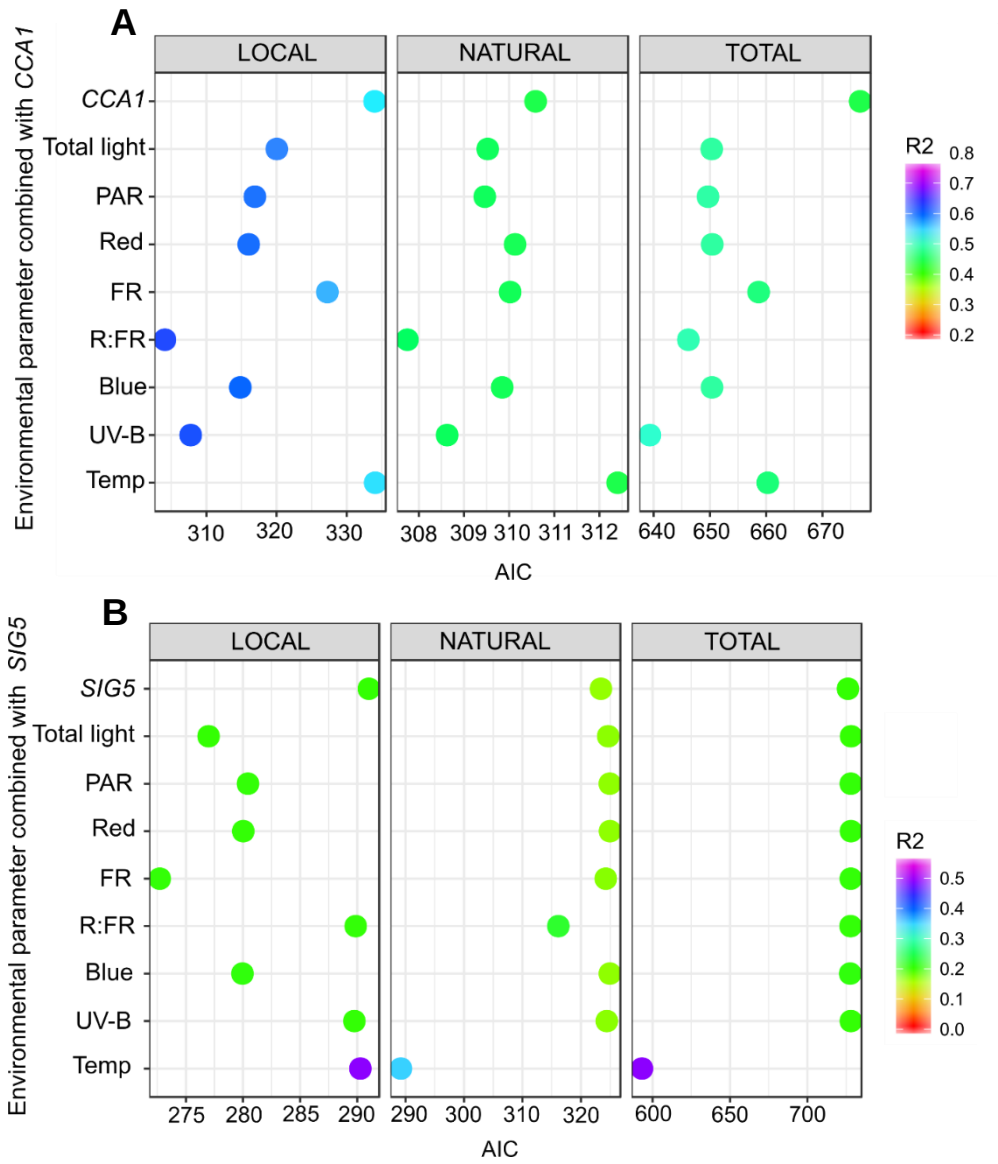


Figure 5-24. Environmental contribution to the model by R:FR is not due to day-night shifts. In *CCA1-SIG5* (local dataset), R:FR remains the environmental cue that improves the model (A). In *SIG5-psbD* BLRP, removal of the dark period results in a shift from temperature to Far Red as main driver of the gene dynamics detected by AIC (local dataset), however, the R^2 is only improved by Temperature (B).

For the *SIG5-psbD* BLRP model temperature remained the major environmental stimulus in the natural and total dataset, however, in the locally manipulated plants FR light was the major environmental cue instead of temperature determined by AIC (Fig.

5-24B), however, the R^2 value is not improved by FR, only by temperature. Also, I wondered why Far Red light specifically, since the correlation plots showed a high correlation with total light intensity and therefore PAR, blue and red light (Table 5-2).

By plotting the environmental cues FR and red light (both are part of the R:FR cue), it was found that the last part of the pattern corresponding to the ND filter treatment had a greater intensity peak for FR but not for red light (Fig. 5-25).

These results indicate that R:FR information is a true environmental cue that regulates the SIG5 pathway independent from day and night cycles and that is the major driver in the dynamics between *CCA1-SIG5* where environmental conditions were manipulated in the field. Furthermore, the results for the *SIG5-psbD* BLRP model showed that FR light also contributes to the signalling transduction between nucleus and chloroplast in the light period.

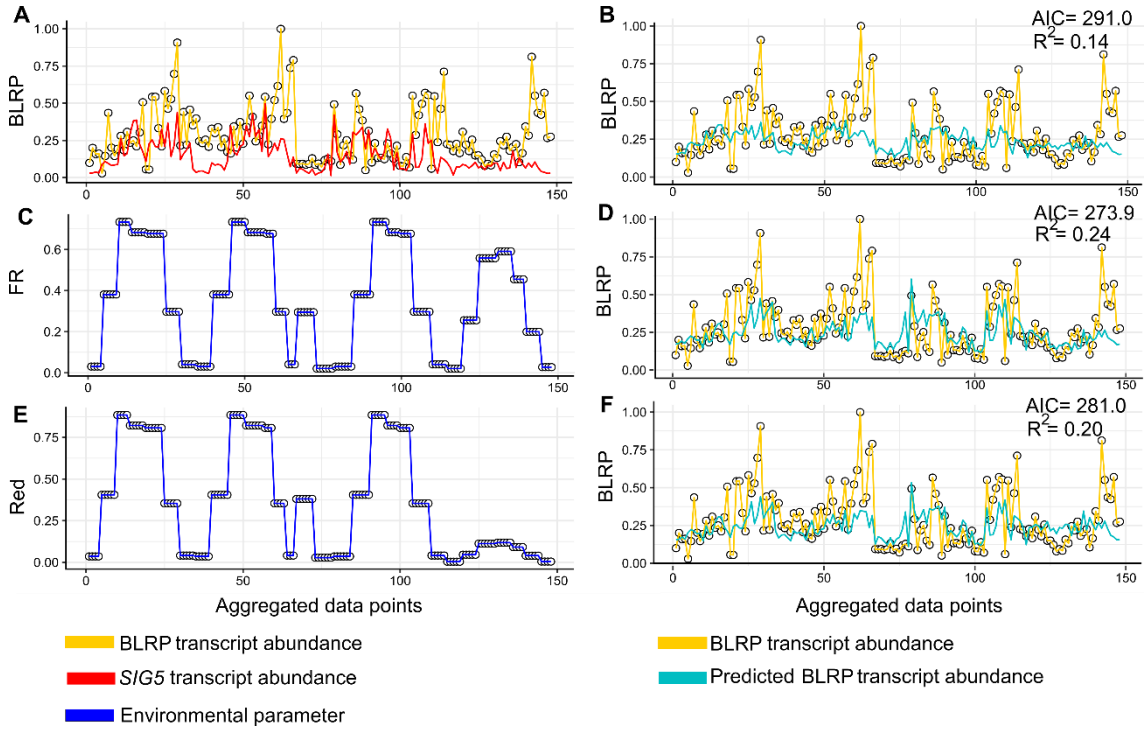


Figure 5-25. In the absence of nocturnal effects, *psbD* BLRP prediction is improved by FR but not by Red light (local dataset). The *SIG5-psbD* BLRP model (A, B), and added environmental term FR (C, D), or Red (E, F).

5.2.10 A consensus dynamic model for SIG5-mediated environmental signal integration

So far, we have determined which are the major environmental stimuli of the two parts of this signalling pathway, however a certain proportion of the dynamics remains unexplained. Also, considering the environmental drivers of each system, it seems that light cues explain most of the experimental data in *CCA1-SIG5* models, whereas temperature is the major environmental cue in *SIG5-psbD* BLRP models.

Therefore, we consider the possibility that the *CCA1-SIG5-psbD* BLRP system could be explained by two environmental parameters rather than just one. To do this, I made combinations of the principal environmental cues and a secondary non-correlated cue, for example UV-B could be paired with R:FR, blue or temperature, and tested them in the models for the entire dataset.

This analysis demonstrated that *CCA1-SIG5* and *SIG5-psbD* BLRP in the total dataset have blue light in common as second environmental cue (Fig. 5-26A).

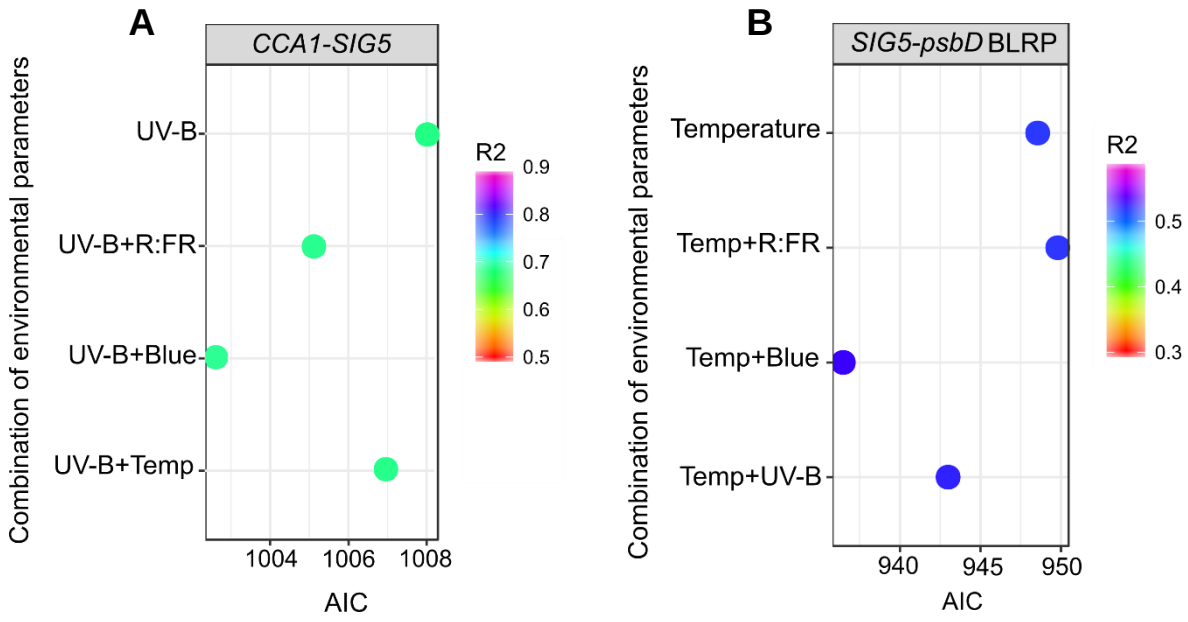


Figure 5-26. Blue light links the three components of the signalling pathway under natural conditions. Analysis of the total datasets indicate that UV-B and blue light in *CCA1-SIG5* (A) and temperature and blue light are the major environmental drivers for *SIG5-psbD* BLRP (B).

Finally, the model equations have the following coefficients the terms:

Equation 5 $\sigma_{SIG5} = 0.46_{CCA1} + 2.2\ UVB - 1.34\ Blue\ light$

Equation 6 $\sigma_{\mathrm{BLRP}} = 0.49_{SIG5} + 1.64\ Temperature - 0.49 Blue\ light$

5.3 CONCLUSIONS

In conclusion, the data revealed seasonal differences in the amplitude of gene transcript accumulation in the signalling pathway and the resulting models suggest the following:

1. There is a time delay between the response of one gene and the following that goes from two hours in the integration of environmental cues and the circadian clock to induce *SIG5* and up to 4 hours for this to take the information to the chloroplast where it regulates *psbD* BLRP.
2. A damped circadian rhythm in winter or spring releases or weakens the regulation between the circadian oscillator and *SIG5* by uncoupling them.
3. Local environment manipulations in the field revealed the importance of the R:FR light input and that is independent from day/night transitions.
4. Overall, UV-B, blue light and temperature control the *CCA1-SIG5-psbD* BLRP signal transduction pathway under natural conditions.

5.4 DISCUSSION

UV-B in plants is perceived by the UV RESISTANCE LOCUS8 (UVR8), through a tryptophan-based mechanism that induces dimers of the photoreceptor to monomerise and interact with CONSTITUTIVELY PHOTOMORPHOGENIC 1 (COP1), an E3 ubiquitin ligase (Rizzini et al., 2011). Downstream elements of the signalling pathway include ELONGATED HYPOCOTYL 5 (HY5) and HY5 HOMOLOG (HYH), which are targeted for degradation by COP1. Mechanistically, the interaction of UV-B with its photoreceptor UVR8 sequesters COP1 and stabilises HY5/HYH (Brown et al., 2005). The coefficient for UV-B (+2.2, Eq. 5) indicates that *SIG5* transcript abundance increases as UV-B increases, which is supported by the molecular mechanism.

The second environmental cue determined by our study was blue light. Cryptochromes (CRY) are the blue light receptors in plants, their inactive state is monomeric and upon excitation homodimerize and stabilise HY5/HYH by interacting with COP1 (Wang et al., 2001). CRY1 acts by inhibiting the interaction between COP1 and SUPPRESOR OF PHYA-105 1 (SPA1), whereas CRY2 sequesters COP1 by stimulating the COP1/SPA1

complex (Liu et al., 2016a). It was shown that blue light through light-stable CRY1 plays a major role in SIG5 signalling to chloroplasts (Belbin et al., 2016;Onda et al., 2008) whereas the light-sensitive CRY2 also promotes photoperiodic flowering (Liu et al., 2008). Our findings show that blue light for *CCA1-SIG5* has a coefficient of -1.34 and for *SIG5-psbD* BLRP a value of -0.49. This means that as blue light increases, the signalling pathway transcripts decrease, which is inconsistent with the molecular mechanism. We must remember that these coefficients were obtained after making a two-component model, therefore I examined the coefficients from the single environmental parameters in the total dataset:

Equation 7 $$\sigma_{SIG5} = 0.45_{CCA1} + 0.82\ Blue\ light$$

Equation 8 $$\sigma_{\mathrm{BLRP}} = 0.39_{SIG5} + 0.14\ Blue\ light$$

The coefficients in the single models have a positive sign (Eq. 7, 8), which means that in the two-model system (Eq. 5,6) there might be correlation between either the environmental terms or between the gene and an environmental term which is causing the model to outweigh for this interaction (Kennedy, 2005). Therefore, we can say that blue light remains the linking environmental clue between the two parts of the signalling pathway, however, the coefficient is not biological significant.

As shown in Chapter 4, HY5/HYH are upstream transcription factors of the response of SIG5 to low temperature, which was the main environmental driver between *SIG5* and *psbD* BLRP identified by field data analysis. However, our model failed to predict half of the variability of *psbD* BLRP, which means there are temperature independent processes that account for that variability. Other explanation for the poor fitting of the model is that chloroplast gene analysis is very noisy, and failure to predict the data comes from technical variability.

Another interesting finding in this chapter was that differences in the magnitude of transcript abundance depend on seasonality, which replicate previous studies (Aikawa et al., 2010). However, the novelty of our work is that these seasonal differences can affect the sensitivity of a signalling transduction pathway, and perhaps the same mechanism operates in other pathways.

Finally, R:FR as an environmental cue was shown to be important when analysing the local environmental manipulation dataset. The biggest change in transcriptional patterns and magnitude comes from the ND filter treatment (Fig. 5-6, 5-7 and 5-8). When we compare the experimental data from the top three environmental cues that improve the

models, we see that the ND filter damped most of them, except for FR which increased. Therefore, R:FR emerged as a primary cue caused by either i) lack of UV-B stimuli because it approximated zero under the ND filter, ii) higher FR light went throughout the filter which altered the R:FR or iii) the effect of both. Previous studies have shown that constant FR alters the circadian rhythm by decreasing transcript accumulation of genes from the morning loop, which includes *CCA1* (Wenden et al., 2011). Nevertheless, the damping of *CCA1* oscillations caused by the ND filter is an interesting tool in the field that could be used to further explain the data from March. As we have previously detailed, during March *CCA1* oscillations are damped and uncoupled from *SIG5* (Fig. 5-15 and 5-16), and we hypothesised this could lead to a basal sensitivity to adverse environmental conditions rather than gated by the circadian clock. The ND filter could then be used to simulate a “winter” clock and we could test this hypothesis in the field.

6 TO FLOAT OR NOT TO FLOAT

"Yes, we all float down here"

6.1 BACKGROUND

Plants need light to fuel photosynthesis, however the availability of light is restricted in submerged environments where the water filters the light. Underwater, low energy light is the first to be filtered (far-red and red), followed by orange, yellow and violet (Talarico and Maranzana, 2000). An important unanswered question is of whether flotation (buoyancy) of aquatic plants enhances their photosynthetic capacity and therefore fitness of these plants under limited light at the bottom and very high light intensities in the surface.

Reports on free floating aquatic plants mention that there are regular appearances of floating plants during certain months every year, while they remain at the bottom in other seasons or cloudy days (Acton, 1916;Harpenslager et al., 2015). These daily and seasonal responses suggest that the circadian clock could have a role in the regulation of these buoyancy responses.

Lake balls or marimo are spherical aggregations of the green macroalgae *Aegagropila linnaei*. Globally extremely rare, they are found in freshwater lakes in the northern hemisphere (Boedeker et al., 2010). Marimo can grow as spherical aggregations, mats, free floating filaments and attached to rocks (Soejima et al., 2008). Unfortunately, marimo balls are endangered, with 50% of the global population reported from herbaria thought to be extinct (Boedeker et al., 2010). The largest populations are reported in Lake Mývatn, Iceland and Lake Akan, Japan. In the latter, they are designated as a natural treasure and are protected by law (Togashi et al., 2014).

Daily and seasonal rhythms of buoyancy in plants are processes very poorly understood. I wished to investigate the processes that cause sinking or floating, and the contribution of circadian regulation to this.

6.2 RESULTS

Using commercially available marimo balls that are believed to originate from Ukrainian populations, we first performed a general morphological characterisation. Fig.6-1A shows the balls in tap water, along with a transverse section (Fig. 6-1B) showing internal cavity. Structurally the balls are formed from tightly tangled filaments (Fig. 6-1C) that are long, thin (Fig. 6-1D) and interlocked by sub-terminal branches (Fig. 6-1E).

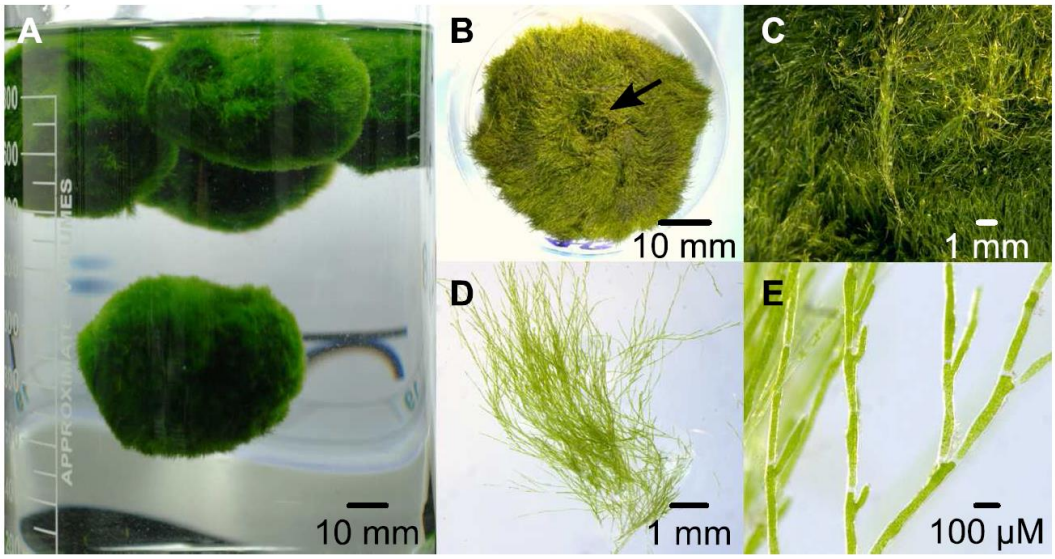


Figure 6-1. Morphology of (A) the ball form of *Aegagropila linnaei*; (B) Transverse section of a ball showing internal cavity (indicated by arrow); (C) Structure consists of filaments tightly tangled in midsection; (D) Shoots in filaments are long and thin interlocked by the branches; (E) Individual filaments showing subterminal insertion of branches. Representative pictures are shown (n= 6).

To investigate the processes that cause sinking or buoyancy, marimo balls were then exposed to different light regimes and quality. It was found that under light/dark (LD) cycles, the balls ascended to the top of a graduated cylinder in the light period and dropped during the dark phase (Fig. 6-2A, 6-2B). The balls were also buoyant under a mix of monochromatic blue and red light, but not green light. We noticed during these experiments that buoyant balls were always covered in small bubbles (Fig. 6-3A), which we reasoned might be the product of photosynthetic activity. To test whether photosynthesis drives buoyancy, a photosynthetic electron transport inhibitor (DCMU, 20µM concentration) was added to the water in the cylinder and buoyancy was recorded. In the presence of DCMU, scarce bubbles appeared in the surface of the ball (Fig. 6-3B) and buoyancy was completely prevented (Fig. 6-3C, 6-3D).

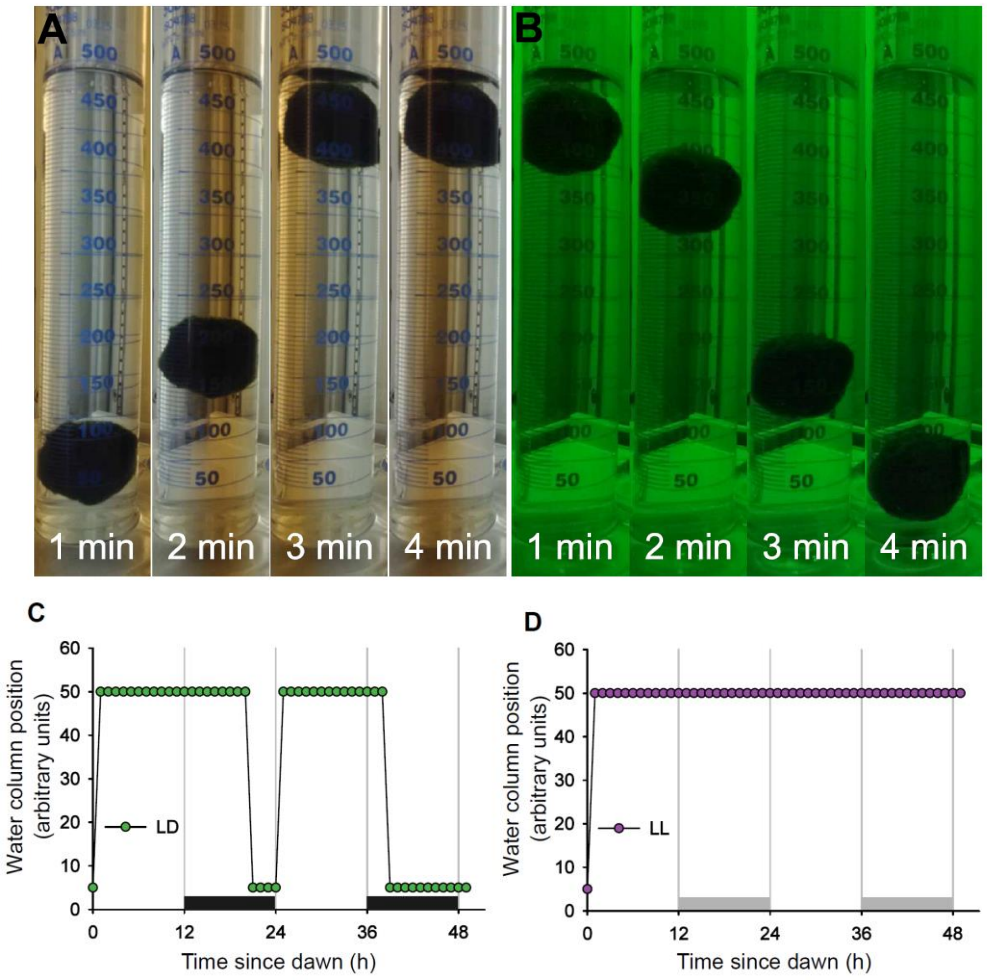


Figure 6-2. Position within a body of water in the light period (A) and dark period (B) changes in minutes under light/dark cycles (12 hours light/12 hours dark). Position of marimo balls within water column under (C) light/dark cycles and (D)constant white light. Representative pictures are shown.

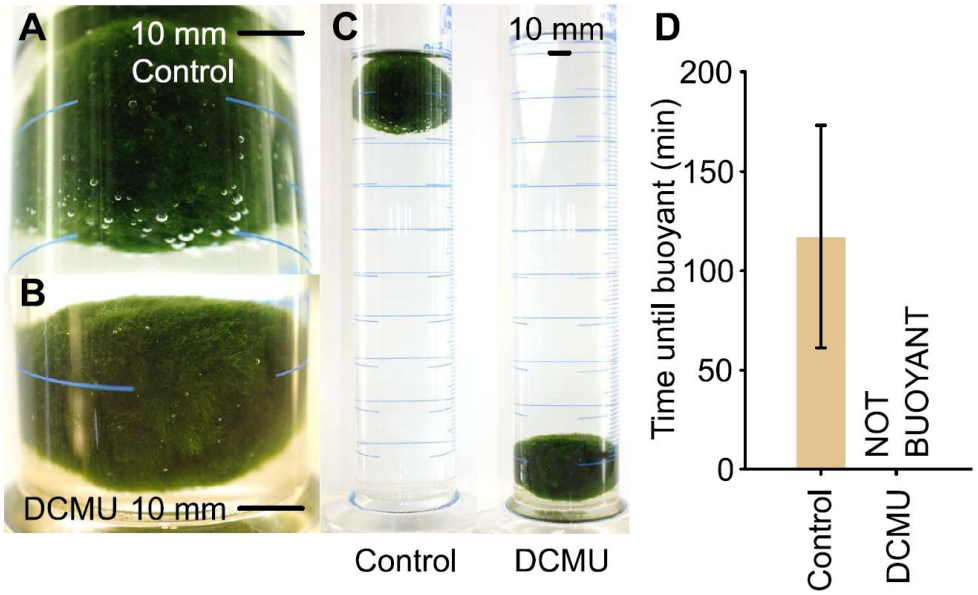


Figure 6-3. Photosynthesis drives buoyancy of marimo. In control condition there are abundant and larger bubbles around the ball (A), while in the presence of a Photosystem II inhibitor there was no buoyancy and scarce bubbles (B, C, and D). Data are mean ± SEM, n=6 balls.

Because there are circadian rhythms in the photosynthetic responses to light of higher plants, we tested whether buoyancy might be circadian regulated (Litthauer et al., 2015). Marimo balls were entrained to 12h LD cycles and then placed under dim constant red light for 24 h to maintain circadian rhythmicity (Noordally et al., 2013), then the balls were exposed to a pulse of blue and red light in the morning (ZT 25) or afternoon (ZT 31).

Results in Fig. 6-4A show that buoyancy takes 81 ± 10 min when the light pulse is given in the subjective morning, whereas in the subjective afternoon, it takes them double the time (178 ± 30 min) suggesting buoyancy is under the control of the circadian clock. To confirm this, marimo balls were placed under dim constant red light for 48 h and then exposed to the blue light pulse. Results show that the marimo balls rise faster during the subjective morning (ZT 49) than in the subjective afternoon (ZT 55). This demonstrates rhythmicity of the buoyancy response across the second and third cycles of constant light.

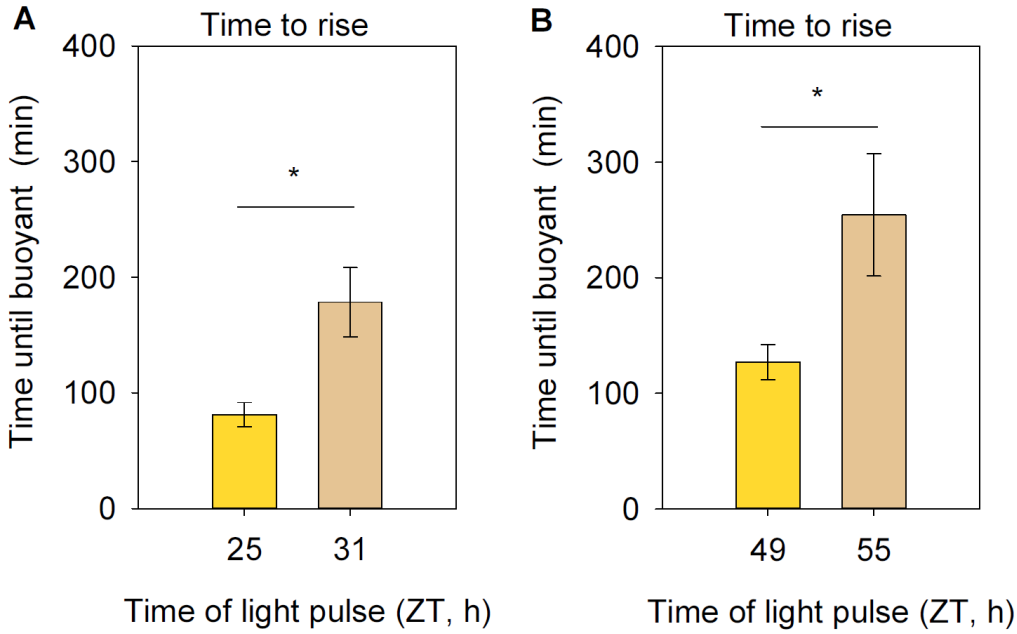


Figure 6-4. Buoyancy under a blue/red light mixture is circadian-gated. (A) After 24 h and (B) 48 hours under constant red light (5 µmol/m^{-2}/s^{-1}) buoyancy was recorded in the presence of red and blue light (20 µmol/m^{-2}/s^{-1}) at the beginning of the subjective day (ZT 25, ZT 49) or halfway the subjective day (ZT 31, ZT 55). Data are mean ± SEM, n=4 balls. * represents statistically significant differences of at least $p<0.05$ using t-student test.

Next, we wished to investigate whether there were differences between day and night in photosynthesis capacity depending on the ball region. Using the chlorophyll fluorescence technique and a narrow fibre optic probe that could be inserted in the interior, mid-point and surface of the ball, the effective photochemical yield of PSII (Y(II)) and the rate of photosynthetic electron transport (ETR) were measured in a range of light intensities using a rapid response curve protocol. Results in Fig. 6-5 show that Y(II) was greater at the marimo surface when the balls were in the light period of the cycle rather than during the dark period, both at low and high light intensities. This day-night photosynthetic rate difference was absent within the marimo interior and mid-point, similarly to the surface during the dark period (Fig. 6-5). Due to this, we concluded that there are anatomically dependent day-night fluctuations on the photosynthetic capacity of the marimo.

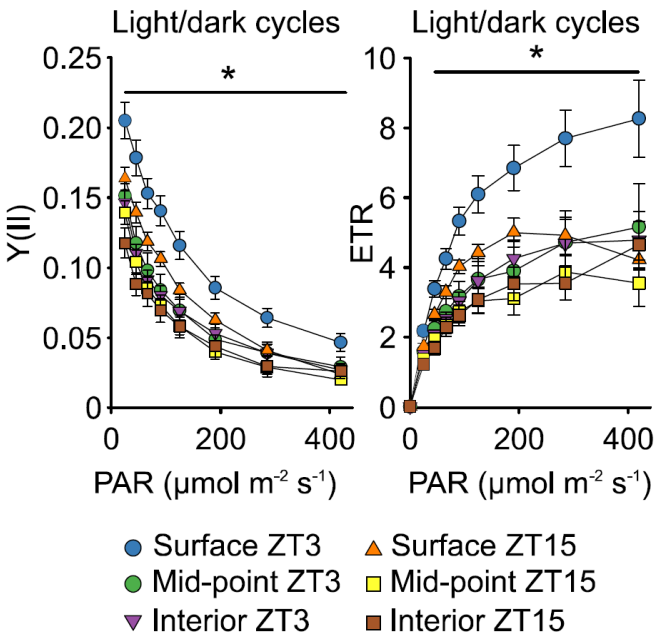


Figure 6-5. Chlorophyll fluorescence parameters Y(II) and ETR from the outside, halfway out and centre locations at day (ZT 3) and night (ZT 15) under 12 h LD cycles. Data are mean ± SEM, n=10 balls. * represent statistically significant differences of at least $p<0.05$ using One Way ANOVA (Student-Newman-Keuls Method).

To detect circadian parameters, timecourses of rapid light response curves were obtained from the surface and interior of the balls, previously entrained to 12 h LD cycles were then placed under constant blue light for 90 min, after this, quantum yield of PSII Y(II) and the non-photochemical quenching (Y(NPQ)) were obtained at many light intensities within the light response curves. This protocol was repeated for 96 hours using a MAXI-PAM imaging system (Walz GmbH, Effeltrich, Germany). The data were analysed using the Fast Fourier Transform Non-Linear Least Squares algorithm (FFT-NLLS) and an independent mathematical method called Maximum Entropy Spectral Analysis algorithm (MESA) available at biodare2.ed.ac.uk (Zielinski et al., 2014). Results in Table 1 summarize the analysis of rhythmicity for Y(II) and Y(NPQ) at every light intensity within the rapid light response curves taken every 2 hours for 5 days in continuous light. In general, the number of rhythmic areas differed between light intensities and between the interior and surface of the balls (maximum 10 and minimum 1 of total areas measured).

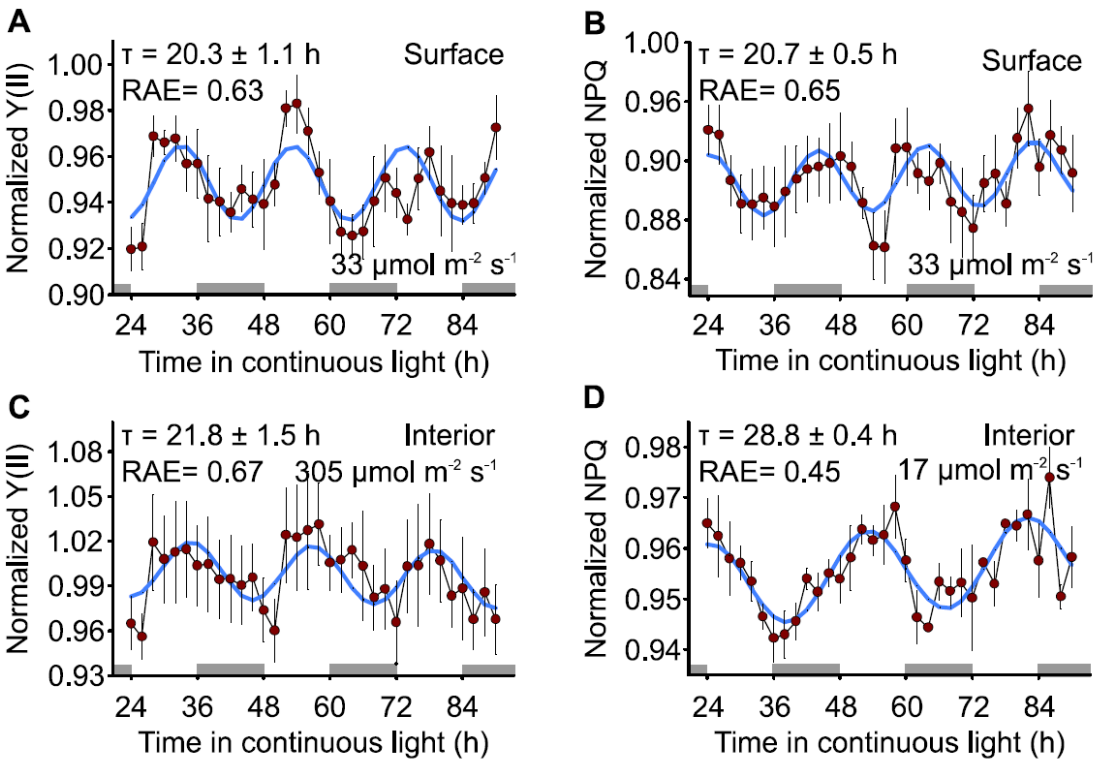


Figure 6-6. Marimo under continuous light were subjected to rapid light response curves (darkness to 425 µmol m^{-2} s^{-1} light intensities) every two hours; Mean period and RAE estimation (FFT NLLS) are shown for the outside and inside of the balls at light intensities with lowest RAE values which corresponded to 33 µmol m^{-2} s^{-1} pulses of blue light in Y(II) (A) and Y(NPQ) (B) for the surface and 305 and 17 µmol m^{-2} s^{-1} in Y(II) (C) and Y(NPQ) (D), respectively, for the interior. Data are mean ± SEM in red circles and blue lines represent FFT-NLLS waveform fit to data; n= 5 balls surface, n=4 balls interior.

In Fig. 6-6 are shown the timecourses with the best fit to a wave model (in blue) measured as RAE with values closer to zero indicate best fit. For the surface, lower RAE by both methods (0.63-0.65 by FFT-NLLS and 0.62-0.72 by MESA) occurred at 33 µmol m^{-2} s^{-1} light intensity for both Y(II) and Y(NPQ), with circadian period estimate of 20.3 +/- 1.1 h and 20.7 +/- 0.5 h, respectively (Fig. 6-6A, 6-6B). However, for the interior, the best fit was at different light intensities for the two photosynthetic parameters within the light curve, 305 and 17 µmol m^{-2} s^{-1} for Y(II) (Fig. 6-6C) and Y(NPQ) (Fig. 6-6D), respectively, and period estimates were too different (21.8 +/- 1.5h and 28.8 +/- 0.4h) to get an accurate estimation. Due to this, delayed fluorescence timecourses, using an EM-CCD camera (Photek Ltd, St Leonardson Sea, UK), were performed in the surface (Fig. 6-7A) and interior (Fig. 6-7B), which resulted in period estimates of 26.0 +/- 0.8 h and 23.9 +/- 1.2 h. Taken together, results suggest that circadian rhythms were more robust in the surface than interior of marimo.

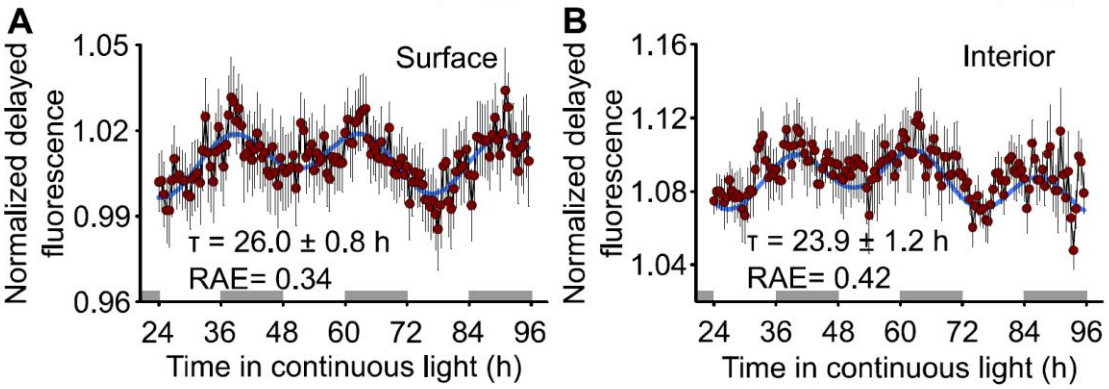


Figure 6-7. Using delayed fluorescence, rhythms in the surface (A) and interior (B) were detected. Data are mean ± SEM in red circles and blue lines represent FFT-NLLS waveform fit to data; n= 3 balls surface, n=3 balls interior.

Table 6-1. Y(II) and Y(NPQ) mean values obtained from light response curves conducted every two hours in the surface and interior areas of *Aegagropila linnaei* balls under continuous blue light. Datum were analysed using the online software available at BioDare 2.0 (https://biodare2.ed.ac.uk). Amplitude and baseline detrending was used for both methods. N is the number of total areas collected (four per subject) and n is number of rhythmic areas.

FFT NLLS								
Parameter	PAR (µmol m^{-2} s^{-1})	N	n	Region	Period	SEM	RAE	SEM
Y(II)	0	20	6	Surface	28.17	0.40	0.64	0.05
Y(II)	1	20	10	Surface	28.36	1.14	0.66	0.04
Y(II)	17	20	6	Surface	26.38	0.63	0.64	0.06
Y(II)	33	20	4	Surface	20.25	1.05	0.63	0.07
Y(II)	73	20	2	Surface	22.85	0.85	0.6	0.14
Y(II)	108	20	3	Surface	25.86	1.73	0.72	0.08
MESA								
Parameter	PAR (µmol m^{-2} s^{-1})	N	n	Region	Period	SEM	RAE	SEM
Y(II)	0	20	6	Surface	27.48	1.03	0.86	0.05
Y(II)	1	20	10	Surface	25.92	1.09	0.9	0.06
Y(II)	17	20	6	Surface	25.68	0.87	0.92	0.06
Y(II)	33	20	4	Surface	20.8	0.78	0.62	0.04
Y(II)	73	20	2	Surface	23.68	0.19	0.68	0.06
Y(II)	108	20	3	Surface	25.64	1.34	0.78	0.03
FFT NLLS								
Parameter	PAR (µmol m^{-2} s^{-1})	N	n	Region	Period	SEM	RAE	SEM
Y(II)	0	16	3	Interior	21.36	2.18	0.98	0.03
Y(II)	17	16	6	Interior	21.78	1.53	0.67	0.02
Y(II)	33	16	2	Interior	21.42	1.14	0.78	0.03

Y(II)	47	16	4	Interior	22.15	0.62	0.83	0.03
Y(II)	73	16	2	Interior	22.87	2.35	0.87	0.13
Y(II)	108	16	2	Interior	20.03	0.87	0.86	0.09
MESA								
Parameter	PAR (µmol m^{-2} s^{-1})	N	n	Region	Period	SEM	RAE	SEM
Y(II)	0	16	3	Interior	22.56	2.68	0.82	0.09
Y(II)	17	16	6	Interior	22.5	1.38	0.62	0.07
Y(II)	33	16	2	Interior	21.38	0.55	0.79	0.04
Y(II)	47	16	4	Interior	22.06	0.71	0.69	0.06
Y(II)	73	16	2	Interior	23.6	3.41	0.9	0.06
Y(II)	108	16	2	Interior	21.22	1.57	1	0.01
FFT NLLS								
Parameter	PAR (µmol m^{-2} s^{-1})	N	n	Region	Period	SEM	RAE	SEM
Y(NPQ)	17	20	3	Surface	19.12	3.39	0.82	0.05
Y(NPQ)	33	20	4	Surface	19.71	0.45	0.65	0.04
Y(NPQ)	47	20	3	Surface	18.8	1.56	0.87	0.06
Y(NPQ)	108	20	3	Surface	11.38	0.98	0.9	0.05
Y(NPQ)	150	20	2	Surface	24.38	2.83	0.97	0.01
Y(NPQ)	195	20	1	Surface	27.47	3.08	0.7	0.07
Y(NPQ)	245	20	6	Surface	24.35	1.12	0.92	0.01
Y(NPQ)	305	20	4	Surface	27.25	1.43	0.8	0.05
MESA								
Parameter	PAR (µmol m^{-2} s^{-1})	N	n	Region	Period	SEM	RAE	SEM
Y(NPQ)	17	20	3	Surface	18.14	3.79	0.69	0.03
Y(NPQ)	33	20	4	Surface	18.86	0.51	0.72	0.03
Y(NPQ)	47	20	4	Surface	17.8	1.68	0.67	0.02
Y(NPQ)	108	20	4	Surface	19.38	0.74	0.79	0.04
Y(NPQ)	150	20	2	Surface	25.42	2.88	0.78	0.03
Y(NPQ)	195	20	3	Surface	29.18	2.73	0.65	0.11
Y(NPQ)	245	20	6	Surface	27.58	1.25	0.76	0.05
Y(NPQ)	305	20	4	Surface	27.08	1.48	0.68	0.02
Y(NPQ)	369	20	6	Surface	27.52	1.74	0.83	0.05
FFT NLLS								
Parameter	PAR (µmol m^{-2} s^{-1})	N	n	Region	Period	SEM	RAE	SEM
Y(NPQ)	17	16	2	Interior	18.18	4.19	0.89	0.03
Y(NPQ)	33	16	3	Interior	29.84	2.93	0.96	0.01
Y(NPQ)	47	16	5	Interior	18.94	2.46	0.85	0.06
Y(NPQ)	73	16	4	Interior	30.79	2.68	0.53	0.03
Y(NPQ)	108	16	2	Interior	31.51	4.70	0.78	0.05
Y(NPQ)	150	16	5	Interior	27.91	0.41	0.64	0.03
Y(NPQ)	195	16	9	Interior	29.17	1.51	0.49	0.06
Y(NPQ)	245	16	6	Interior	28.82	0.27	0.54	0.05
Y(NPQ)	305	16	3	Interior	28.76	0.35	0.45	0.04
Y(NPQ)	369	16	4	Interior	29.04	0.84	0.62	0.02

MESA								
Parameter	PAR (µmol m^{-2} s^{-1})	N	n	Region	Period	SEM	RAE	SEM
Y(NPQ)	17	16	2	Interior	19.34	2.52	0.8	0.05
Y(NPQ)	33	16	3	Interior	27.54	2.92	0.76	0.06
Y(NPQ)	47	16	5	Interior	29.54	2.02	0.59	0.06
Y(NPQ)	73	16	4	Interior	29.38	2.72	0.66	0.06
Y(NPQ)	108	16	2	Interior	26.36	2.20	1	0.09
Y(NPQ)	150	16	5	Interior	30.2	1.17	0.95	0.06
Y(NPQ)	195	16	9	Interior	29.32	1.60	0.59	0.04
Y(NPQ)	245	16	6	Interior	30.2	0.47	0.69	0.04
Y(NPQ)	305	16	3	Interior	28.82	0.53	0.56	0.06
Y(NPQ)	369	16	4	Interior	28.66	0.69	0.75	0.03

In conclusion, we identified diurnal and circadian fluctuations in the buoyancy response to light of marimo and found that this buoyancy is due to photosynthesis. Also, we identified circadian rhythms of two measures of photosynthesis in this filamentous freshwater alga.

6.3 DISCUSSION

Photosynthetic organisms must take carbon from the environment. For land plants the source is gaseous carbon dioxide (CO_2), which is a molecule that can easily cross membranes. However, aquatic plants must overcome the fact that CO_2 dissolves in water as bicarbonate (HCO_3^-), where it has a 10^4 slower diffusion rate compared to air (Gutknecht et al., 1977). Besides, the low affinity of Rubisco for carbon dioxide forced aquatic photosynthetic organisms to develop carbon concentrating mechanisms (CCM) (Giordano et al., 2005). In eukaryotic algae, the CCM consist of a structure inside the chloroplast called pyrenoid which are largely composed of Rubisco (Borkhsenious et al., 1998).

Bicarbonate, unlike CO_2, is not cell membrane permeable and needs specialised transporters (carbonic anhydrases) which also function as part of the CCM by stopping the carbon inside the cell to diffuse back to the environment (Giordano et al., 2005). Carbonic anhydrases are also found in pyrenoids and catalyse the conversion of HCO_3^- back to CO_2, substrate for Rubisco. Nevertheless, carbonic anhydrases require energy for their activity as ATP, NADH, ferredoxin or coupled to a sodium gradient. Therefore, carbon transport is an active process that needs to be coordinated with photosynthesis (Moroney et al., 2001).

In this study we report that *Aegagropila linnaei* buoyancy is due to photosynthesis and this is controlled by the circadian clock.

6.3.1 Further experiments

An interesting idea that I would like to test in future is whether floating also allows marimo to reach the surface of water and take carbon dioxide from the air. Therefore, the circadian rhythm of photosynthesis could be coordinating buoyancy and carbon dioxide uptake. The advantages of this would be to have access to higher concentrations of carbon dioxide, which in air is not energetically expensive to uptake, at critical times when photosynthesis produces the ATP necessary to store it in pyrenoids, which are found in marimo (Boedeker, 2010).

This could be tested by placing marimo under continuous blue light as in previous experiments (Fig. 6-5) but stopping a group of them to reach the surface with a barrier (mesh) and then measuring growth, starch accumulation or looking at the pyrenoids. We would expect that stopping marimos to float to the surface when they have reached buoyancy would cause slow growth and lower carbohydrate biosynthesis.

Reports on free floating aquatic plants mention that there are regular appearances of floating plants during certain months every year, while they remain at the bottom in other seasons or cloudy days (Acton, 1916;Harpenslager et al., 2015). These daily and seasonal responses suggest that the circadian clock could have a role in the regulation of these buoyancy responses throughout the year.

Temperatures in two of the lakes which harbour the biggest marimo populations can reach below zero depending on time of year. Temperature in lake Akan in Japan goes from -6 °C to 15 °C while in lake Mývatn in Iceland it goes from -0.2°C to 10.2°C in the warmest month, July.

For this study, marimo was kept at 15 °C which was not the average temperature in its natural habitat, therefore I would like to study the extent to which the circadian rhythm of photosynthesis and buoyancy are affected by temperatures in different seasons. Furthermore, we could investigate whether the circadian clock of marimo can also give an estimate of daylength by putting it under different photoperiods and study the effect upon photosynthesis and buoyancy.

Furthermore, it would be informative to obtain the genome of marimo and investigate which clock components are present and the overall circadian clock architecture.

6.4 ACKNOWLEDGEMENTS

Miss Tara Saskia de Fraine and Miss Olivia Griffiths, third year project students from the Biology degree at University of Bristol who I supervised also collected data for Fig. 6-1B, 6-1C and Fig. 6-5.

6.5 NOTE

This research is published as:

CANO-RAMIREZ, D. L., SASKIA DE FRAINE, T., GRIFFITHS, O. G. & DODD, A. N. 2018. Photosynthesis and circadian rhythms regulate the buoyancy of marimo lake balls. *Current Biology*, 28, R869-R870.

7 GENERAL DISCUSSION

7.1 SIG5-DEPENDENT SIGNALLING PATHWAY

Laboratory data indicates that LT induces both *SIG5* and *psbD* BLRP transcription, in a circadian gated manner (Fig.4-3C, 4-3D), which is consistent with the expression under natural conditions (Fig. 5-4). However, *psbD* BLRP transcript abundance was not greatly induced in the field by the high levels of *SIG5* during the night (Fig. 5-4C, 5-4D), which could point to the role of the clock in suppressing the pathway at night. Supporting this, *SIG5* induction in response to low temperature is a light-independent process (Fig. 4-6C) but *psbD* BLRP induction is light dependent (Fig. 4-6D). Therefore, we can infer that SIG5 activity in the chloroplast is controlled by light either during import or modulating its activity through post translational modifications. To gain insights into the potential for post-translational regulation, I used homology modelling. Homology modelling of SIG5 and predicted phosphorylation sites mapping in the structure (Fig. 3-15A) revealed that SIG5 has a putative phosphorylation site in the area where SIG1 has been shown to be regulated by phosphorylation in response to redox imbalance (Shimizu et al., 2010). However, the position in the case of SIG5 overlapped with what might be the transit peptide and therefore could instead be involved in the regulation of SIG5 import (May and Soll, 2000). To solve this, firstly, I would need to verify that the putative phosphorylation site is phosphorylated *in vivo*. Secondly, I would mutate the phosphorylation site and monitor *psbD* BLRP transcription under light dark cycles, continuous light and continuous darkness. Finally, I would like to know whether import of SIG5 happens in the dark. To do this there are two possibilities 1) isolate chloroplast, extract proteins and check by immunoprecipitation whether SIG5 is effectively imported or 2) fuse SIG5 to a fluorescent probe small enough to allow importation into chloroplasts and use confocal microscopy.

It is important to note that in this work I investigated signalling pathway regulation mostly through transcript abundance measured as qPCR-RT because it has been shown that 56% to 84% of protein variation can be explained by mRNA abundance whilst translation rate explains only 9% of the protein abundance variability (Li et al., 2014) when protein and mRNA levels remain stable over time (several hours).

In the case of dynamic adaptation processes (stress responses, cell differentiation, etc.), nearly 80% of alterations in protein abundance are explained by altered mRNA levels in the case of yeast fission (Lackner et al., 2012). However, it is necessary to consider the

temporal delay between transcript induction and protein abundance because maturation, export, and translation of mRNA takes time (Liu et al., 2016b). Furthermore, this delay was found to be protein-specific in different systems. In circadian-regulated processes, protein levels followed mRNA oscillations with a delay of about 6h and the length of this delay varies during the day in mice liver (Robles et al., 2014). In Arabidopsis, it was shown that high translation rates coinciding with high transcript levels contribute to photoperiodic-dependent regulation of proteins and this is controlled by the circadian clock (Seaton et al., 2015). In this work, the delay between *SIG5* transcription and *psbD* BLRP induction (Fig. 4-3 and 5-3) supports these previous studies.

Another level of protein regulation is through post transcriptional processes, which are crucial for short term responses (phosphorylation and ubiquitination) and protein turnover. In Arabidopsis it has been shown that there are oscillations in the phosphorylation of critical regulators of clock components and other regulators of various processes (Choudhary et al., 2015).

Furthermore, protein activity is regulated by spatial compartmentalisation, inhibitors, activators, co factor availabilities and environmental factors (such as pH). However, in the case of transcription factors, it is known that they are translated on demand, where upregulation is achieved through activation of pre-existing transcripts translation (Jovanovic et al., 2015) and incorporated to protein complexes.

7.2 BUT THE WORLD IS COMPLICATED

Here I identified that temperature is one of the main drivers of the SIG5-dependent signalling pathway, with a coefficient value of 1.64 for SIG5 predicting *psbD* BLRP (Eq. 6). However, the sign of the coefficient is positive, meaning the relationship between temperature and the interaction *SIG5-psbD* BLRP is a positive one. This could be interpreted as an increase of one degree in temperature induces almost a two fold change in *psbD* BLRP transcript abundance.

Previously, modelling of 461 microarray data including meteorological data from rice in the field concluded that the main environmental driver of *SIG5* (Os05g0586600) was temperature. However, the study did not cover chloroplast encoded genes so we were unable to find data for *psbD* (Nagano et al., 2012). In the linear model from that study, temperature had a coefficient of -5.9 so the relationship with temperature is inverse (retrieved from https://fitdb.dna.affrc.go.jp/ and accessed 12/09/2018). For example, a drop of one unit of temperature increases the mean transcription of *SIG5* almost 6 units.

From laboratory data detailed in Chapter 4, we know that low temperature increases the transcript abundance of *SIG5* and this in turn induces *psbD* BLRP transcription. This coincides with the study in rice but not with our model, as temperature was not the main environmental component in the prediction of *SIG5* by *CCA1* (Eq. 5). This could be explained by the differences between the conditions studied in Chapter 4 and what happens in the field. In Chapter 4, temperatures of 4 °C were applied to the plants, however such low temperatures were only witnessed during sampling in March (Fig. 5-3E). The model equation obtained from the analysis of natural conditions (March and September only) for *CCA1-SIG5* including temperature are the following:

Equation 9 $$\sigma_{SIG5} = 0.52_{CCA1} - 0.13\, Temperature$$

The coefficient for temperature is negative, which indicates that decreasing temperature increases *SIG5* transcript abundance. However, under natural conditions it seems that light input to the signalling overrides the effect of temperature.

One important difference existed in the environmental data used by Nagano et al. (2012), because light was considered as global solar radiation and not dissected into its components. Another difference was the hypothesis tested in Chapter 4 and the one tested by Nagano et al. (2012), where the transcription dynamic of each gene was independently modelled, considering each gene to be regulated independently, whereas we modelled the dynamic between the genes of the pathway. Our purpose was to reconstruct the dynamic regulation between the components of the signalling pathway and therefore the model for *SIG5* depended on *CCA1* and *psbD* BLRP on *SIG5*.

Another possible explanation comes from the type of plant being studied. Rice cultivars differ in their sensitivity to UV-B light. For example, samples from Nagano et al. (2012) came from two types of rice varieties *Oryza sativa* L. spp. *japonica* cv. Nipponbare or cv. Norin 8 which are UV-resistant and UV-sensitive cultivars respectively (Hidema and Kumagai, 2006). Rice UV-B sensitivity is mostly caused by deficient photorepair and excision repair of DNA damage (Hidema et al., 1997). Therefore, it is possible that in rice UV-B is sensed as a threat rather than an environmental cue and *SIG5* has a more temperature related role than a UV-B signalling one.

Rice is also temperature sensitive which limits its harvesting season and the variety used in crop production is non-perennial. *A. thaliana* is also short-lived and dies after seed setting. However, *A. halleri* is a perennial, which means it withstands conditions that the other plants are never exposed to and is freezing tolerant.

Because most crops of commercial importance are sensitive to temperatures less than 10 ºC, frosts and freezing temperatures cause massive economic losses. For example, in the UK, last year's cold weather lead to a drop by 70-80% of British apples, pears and plums while winemakers report catastrophic crop damage due to frost. Other crops such as spring barley, wheat and sugar beet struggled to develop.

In this light, our studies about freezing tolerance can be applied to temperature sensitive crops such as rice while the seasonal changes of the same signalling pathway can help understand perennial edible crops and tree crops.

7.3 SEASONALITY

During the evolution of plants, the first land colonising plants came from freshwater filamentous or single-celled green algae (Delwiche and Cooper, 2015), which possibly adapted to high light intensities by being buoyant. Therefore, understanding the roles of buoyancy in photosynthetic adaptation will also provide insights into the colonization of the land by plants. An important unanswered question is of whether flotation (buoyancy) of aquatic plants enhances their photosynthetic capacity and therefore fitness of these plants under limited light at the bottom and very high light intensities in the surface.

The circadian clock in plants, besides giving an estimate of time of day within the cell also tracks changes in day length, which depends on the season. More than a hundred years ago Elizabeth Acton described "Cladophora balls" (now known as *Aegagropila*) collected from Loch Kildona, Scotland. There she describes that the balls appear in the surface of the lake during April and May (Acton, 1916). This seasonality was the first indicative for us to infer the possibility that buoyancy could be under the control of the circadian clock (Chapter 6).

The oscillation of most of the clock genes at 4 °C is damped (Bieniawska et al., 2008). How this affects photosynthesis is poorly understood, especially in aquatic plants. We hypothesise that winter would lower the amplitude of the clock and uncouple it from other processes as is the case between *CCA1* and *SIG5* (Chapter 5).

Taking the example from *A. halleri*, we would expect the clock of marimo to damp too during winter due to low temperature and shorter days. Since Lake Akan in Japan, which harbours one of the biggest populations of marimo, is reported to freeze during the winter we would expect that the uncoupling of the circadian rhythm with buoyancy would lead

to marimo to stay in the bottom of the lake. If there was no circadian rhythm and buoyancy depended only on light, a bright winter day would cause marimo to rise to the surface, freeze and get trapped in the ice killing it.

We hypothesise there will be differences in the circadian clock regulation between seasons, because during winter it is unlikely plants will float (even during bright and warmer days) because they can risk being trapped in ice formed in the surface.

7.4 EVOLUTION OF THE TEMPERATURE ENTRAINMENT TO THE CIRCADIAN CLOCK

The strongest circadian entrainment cues are light and temperature. Plants need light to fuel photosynthesis, however the availability of light is restricted in submerged environments where the water filters the light (Vadeboncoeur et al., 2008). Furthermore, water has a higher heat capacity, which means it needs more energy to increase the temperature one degree and therefore temperature fluctuations are less drastic than in terrestrial conditions. In addition to acting as an entrainment cue, circadian clocks of higher plants have a property called temperature compensation, which ensures that the circadian rhythm is robust to fluctuations in the environmental temperature (Eckardt, 2006). Therefore, we hypothesise that underwater plants such as marimo have a less robust temperature compensation system (Fig.7-1).

However, genome analysis of the closest aquatic and land relatives of plants, charophytes and bryophytes, has revealed that most of the core clock genes identified in the model plant *Arabidopsis thaliana* were already present (Linde et al., 2017).

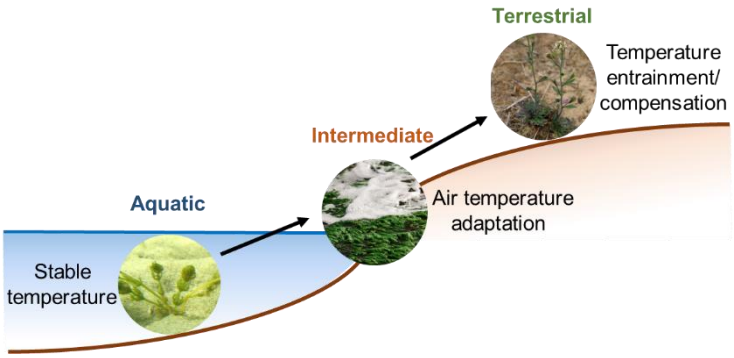


Figure 7-1. Modern closest relatives of land plants are exclusively found in freshwater habitats. We hypothesise that circadian clocks evolved temperature entrainment and compensation mechanisms to adapt to life in land.

Entrainment of the circadian clock by temperature fluctuations and its evolution in plants is a process that is very poorly understood. It might be informative to investigate this process using marimo, to determine: i) whether the clock is entrained by cyclic changes in temperature as it is in Arabidopsis, ii) the range and minimum difference of temperatures by which this applies and, iii) the contribution of the clock genes to this process. The outcomes of this investigation will be key in further understanding the processes that lead to land colonisation by plants.

7.5 GENERAL CONCLUSIONS

This thesis contributes the following novel findings to the field of plant sciences:

a) Three-dimensional structural modelling of *Arabidopsis thaliana* sigma factors which suggests new hypotheses about the evolution of surface charges and regulation of sigma factors by phosphorylation.
b) Identification of a novel cold signalling pathway to the chloroplast and a possible cold-induced retrograde signalling pathway to the nucleus.
c) Identification of the main drivers of the SIG5-dependent signalling to the chloroplast under natural conditions.
d) Characterisation of the buoyancy response of a rare alga and its regulation by circadian rhythms of photosynthesis (Cano-Ramirez et al., 2018).

8 APPENDIX

8.1 APPENDIX A: CORRELATION PLOTS

Table 1. Correlation plots between lagged CCA1 and SIG5 in the local environment conditions.

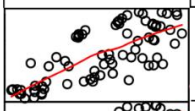
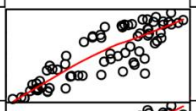
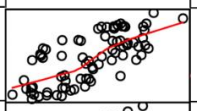
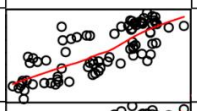
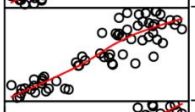
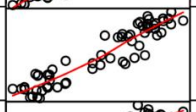
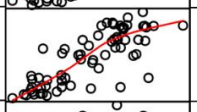
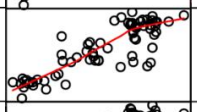
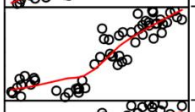
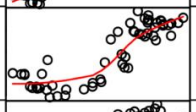
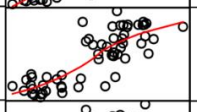
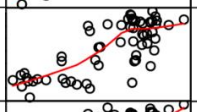
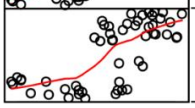
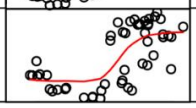
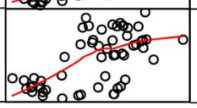
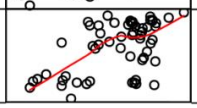
	Control *SIG5*	R^2	Warm *SIG5*	R^2	Chilling *SIG5*	R^2	ND filter *SIG5*	R^2
No lag		0.56		0.74		0.52		0.61
-2h lag		0.81		0.88		0.64		0.72
-4h lag		0.79		0.74		0.67		0.59
-6h lag		0.49		0.49		0.35		0.29

Table 2. Correlation plots between lagged *SIG5* and *psbD* BLRP in the local environment conditions.

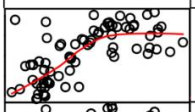
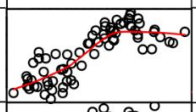
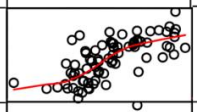
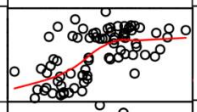
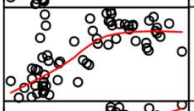
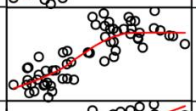
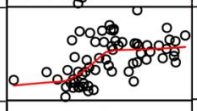
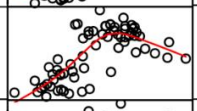
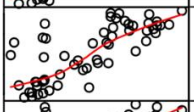
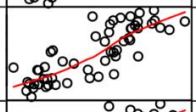
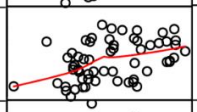
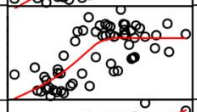
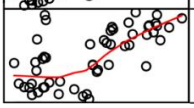
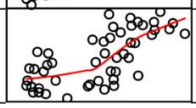
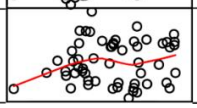
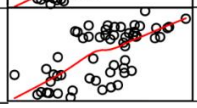
	Control BLRP	R^2	Warm BLRP	R^2	Chilling BLRP	R^2	ND filter BLRP	R^2
No lag		0.45		0.66		0.52		0.37
-2h lag		0.50		0.66		0.27		0.45
-4h lag		0.58		0.69		0.18		0.46
-6h lag		0.42		0.48		0.01		0.44

Table 3. Correlation plots between lagged *CCA1* and *SIG5* for the natural dataset.

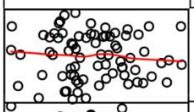
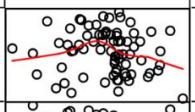
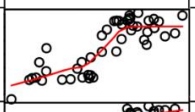
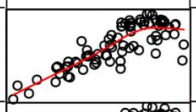
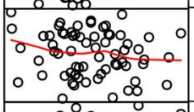
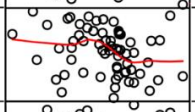
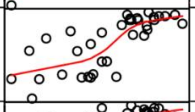
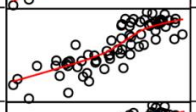
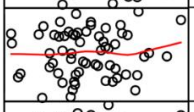
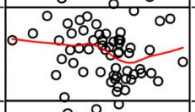
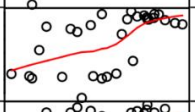
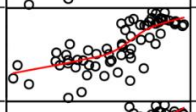
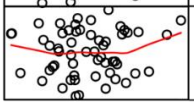
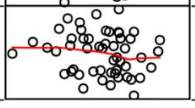
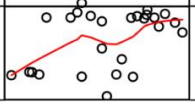
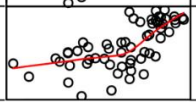
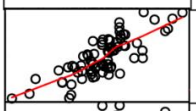
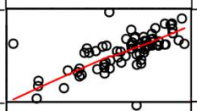
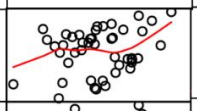
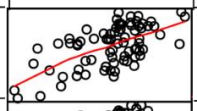
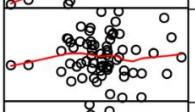
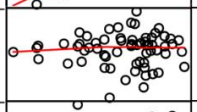
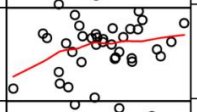
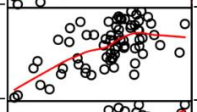
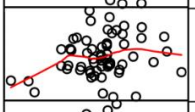
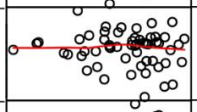
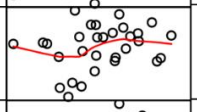
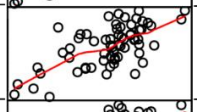
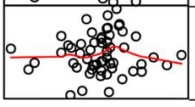
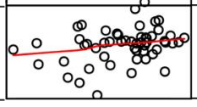
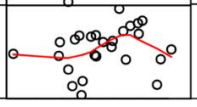
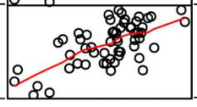
	March sun *SIG5*	R^2	March shade *SIG5*	R^2	September sun *SIG5*	R^2	September sun *SIG5*	R^2
No lag		-0.002		-0.01		0.67		0.66
-2h lag		-0.0002		0.0008		0.56		0.72
-4h lag		-0.005		-0.0004		0.38		0.59
-6h lag		0.004		-0.02		0.22		0.44

Table 4. Correlation plots between lagged *SIG5* and *psbD* BLRP for the natural dataset.

	March sun BLRP	R^2	March shade BLRP	R^2	September sun BLRP	R^2	September sun BLRP	R^2
No lag		0.64		0.50		0.06		0.38
-2h lag		0.0002		0.0002		0.13		0.36
-4h lag		0.04		-0.01		0.06		0.44
-6h lag		0.00002		0.08		0.06		0.44

8.2 APPENDIX B: MODEL SCRIPT

Model script for R

```
#Filter function#

ma <- function(x,n=5)( sides=1))

Ema2h=E
for (i in 1:ncol(Ema2h)) {
    Ema2h[,i] = ma(Ema2h[[i]],n=24)
}

#Model one component: C= SIG5 D=psbD BLRP #

mydata=C
fita=glm((SIG5)~(CCA1),data=mydata, family = gaussian)
fitalight=glm((SIG5)~ (CCA1) +Light,data=mydata, family = gaussian)
fitaPAR=glm((SIG5)~(CCA1)+PAR,data=mydata, family = gaussian)
fitaRed=glm((SIG5)~ (CCA1) +Red,data=mydata, family = gaussian)
fitaFR=glm((SIG5)~ (CCA1) +FR,data=mydata, family = gaussian)
fitaRFR=glm((SIG5)~ (CCA1)+RFR,data=mydata, family = gaussian)
fitaBlue=glm((SIG5)~ (CCA1)+Blue,data=mydata, family = gaussian)
fitaUVB=glm((SIG5)~ (CCA1)+UVB,data=mydata, family = gaussian)
fitaTemp=glm((SIG5)~ (CCA1)+Temp,data=mydata, family = gaussian)

mydata2=D
fitb=glm((BLRP)~(SIG5),data=mydata2, family = gaussian)
fitblight=glm((BLRP)~(SIG5)+Light,data=mydata2, family = gaussian)
fitbPAR=glm((BLRP)~(SIG5)+PAR,data=mydata2, family = gaussian)
fitbRed=glm((BLRP)~ (SIG5) +Red,data=mydata2, family = gaussian)
fitbFR=glm((BLRP)~ (SIG5) +FR,data=mydata2, family = gaussian)
fitbRFR=glm((BLRP)~(SIG5)+RFR,data=mydata2, family = gaussian)
fitbBlue=glm((BLRP)~(SIG5)+Blue,data=mydata2, family = gaussian)
fitbUVB=glm((BLRP)~(SIG5)+UVB,data=mydata2, family = gaussian)
fitbTemp=glm((BLRP)~(SIG5)+Temp,data=mydata2, family = gaussian)

#double environmental parameter#

mydata=C
fita=glm((SIG5)~(CCA1)+E1,data=mydata, family = gaussian)
fitaRFR=glm((SIG5)~ (CCA1)+E1+RFR,data=mydata, family = gaussian)
fitaBlue=glm((SIG5)~ (CCA1)+E1+Blue,data=mydata, family = gaussian)
fitaTemp=glm((SIG5)~ (CCA1)+E1+Temp,data=mydata, family = gaussian)

mydata2=D
fitb=glm((BLRP)~(SIG5)+E1,data=mydata2, family = gaussian)
fitbRFR=glm((BLRP)~(SIG5)+E1+RFR,data=mydata2, family = gaussian)
fitbBlue=glm((BLRP)~(SIG5)+E2+Blue,data=mydata2, family = gaussian)
fitbUVB=glm((BLRP)~(SIG5)+E3+Temp,data=mydata2, family = gaussian)
```

8.3 APPENDIX C: PRIMERS

Table 1. Primers used for qPCR analysis of *A. thaliana*.

Primer	Sequence (5' to 3')
*ACT2*_Forward	TCAGATGCCCAGAAGTGTTGTTCC
*ACT2*_Reverse	CCGTACAGATCCTTCCTGATATCC
psbD BLRP_Forward	GGAAATCCGTCGATATCTCT
psbD BLRP_Reverse	CTCTCTTTCTCTAGGCAGGAAC
*SIG1*_Forward	CCTTGTTCAGGGTGGTCTTATC
*SIG1*_Reverse	CCCTGTCGAATCCACCAATATAC
*SIG2*_Forward	GCGGTTCTACTCTCTCCTAAAC
*SIG2*_Reverse	ACTGCTTCTGGATCTGCTATC
*SIG3*_Forward	GAGGCGTCCATATTCCTTCATC
*SIG3*_Reverse	CTCCTCATCAGCCGCATATTT
*SIG4*_Forward	TCCCTACAATCTCTCCCTTACTC
*SIG4*_Reverse	GCAAACCAGAGCCGTACTAA
*SIG5*_Forward	GTGTTGGAGCTAATAACAGCAGACA
*SIG5*_Reverse	TGTCGAATAACCAGACTCTCTTTCG
*SIG6*_Forward	CGGATGCAGATTCGCTACTT
*SIG6*_Reverse	CTCCGGTAAACGGATTGTTCT
*CBF1*_Forward	CCGCCGTCTGTTCAATGGAATCAT
*CBF1*_Reverse	TCCAAAGCGACACGTCACCATCTC
*CBF2*_Forward	AACTCCGGTAAGTGGGTGTG
*CBF2*_Reverse	CGGCGTATAAATAGCCTCCA
*CBF3*_Forward	ACAGAGGAGTTCGTCGGAGA
*CBF3*_Reverse	ACCAACGTCTCCTCCATGTC
*COR15a*_Forward	GGCCACAAAGAAAGCTTCAG
*COR15a*_Reverse	CTTGTTTGCGGCTTCTTTTC
*SFR2*_Forward	ATGGGCTCTTTGATATCGGG
*SFR2*_Reverse	GCTTTAGTCCAGCACCACAC

Table 2. Primers used for qPCR analysis of *A. halleri.*

Primer	Sequence (5’ to 3’)
*ACT2*_Forward	TCAGATGCCCAGAAGTGTTGTTCC
*ACT2*_Reverse	CCGTACAGATCCTTCCTGATATCC
psbD BLRP_Forward	GGAAATCCGTCGATATCTCT
psbD BLRP_Reverse	CTCTCTTTCTCTAGGCAGGAAC
*SIG5*_Forward	GTGTTGGAGCTAATAACAGCAGACA
*SIG5*_Reverse	TGTCGAATAACCAGACTCTCTTTCG
*CCA1*_Forward	GCACTTTCCGCGAGTTCTTG
*CCA1*_Reverse	TGACTCCTTTCTTATCCTGTTATTCTG

Current Biology
Magazine

CellPress

Correspondence

Photosynthesis and circadian rhythms regulate the buoyancy of marimo lake balls

Dora L. Cano-Ramirez,
Tara Saskia de Fraine,
Olivia G. Griffiths, and Antony N. Dodd[1,*]

Marimo are unusual, attractive and endangered spherical aggregations of the filamentous green macroalga *Aegagropila linnaei* (Figure 1A–E) [1]. Globally rare, marimo populations persist in cold freshwater lakes in Japan, Iceland and Ukraine. Marimo occupy both the lake bed and rise to the lake surface [2,3]. Here, we show that marimo buoyancy is conferred by bubbles arising from photosynthesis. We find that light-induced acquisition of buoyancy by marimo is circadian-regulated. We identify that there are circadian rhythms of photosynthesis in marimo, which might explain the circadian rhythm of buoyancy in response to light. This identifies a circadian-regulated buoyancy response in an intriguing and little-studied plant.

Marimo balls occupy the lake bed and also rise to the surface [2,3]. These movements occur under laboratory conditions, with marimo rising rapidly after acquiring buoyancy (Figure 1A,B). Bubbles form on the surface of buoyant marimo (Figure 1C) [2]. We reasoned that the bubbles might be products of photosynthesis and confer buoyancy, because marimo sink in darkness (Figure 1B) [2]. 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU), which inhibits photosynthetic electron transport from photosystem II (PSII) to plastoquinone, prevented surface bubble formation and marimo buoyancy (Figure 1C–E). DCMU-incubated marimo did not float during 48 h of continuous illumination (Figure 1F). Therefore, the dynamic day–night buoyancy of marimo is due to photosynthesis.

Given that photosynthesis in plants and algae is controlled by the circadian clock [4], we asked whether light-induced acquisition of buoyancy

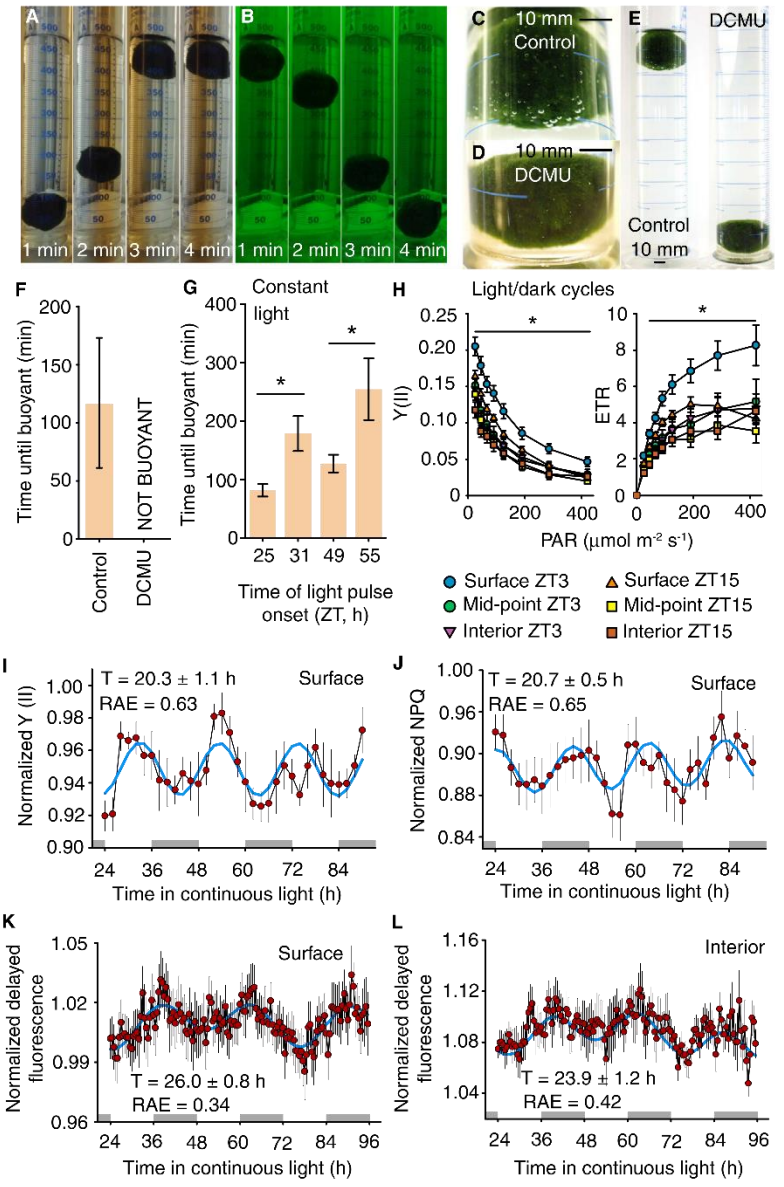


Figure 1. Photosynthesis, circadian rhythms and marimo buoyancy.
(A,B) Marimo position within a water column changes rapidly and is bimodal, with balls (A) buoyant in light and (B) sunken in physiological darkness (time in minutes). (C–E) 20 µM DCMU prevents buoyancy and surface bubble formation. (F) DCMU-treated marimo do not rise within the water column (n=3). (G) Circadian regulation of the acquisition of buoyancy in response to illumination (n=4). (H) Quantum yield of PSII (Y(II)) and electron transport rate (ETR) within rapid light curves (RLCs) from three positions within the marimo ball at ZT3 (daytime) and ZT15 (night), under 12 h light/12 h dark cycles (n=10). The marimo exterior during the day differs significantly from the other measures. (I,J) Circadian oscillations in two photosynthetic parameters on the marimo surface. Data are from a timecourse of RLCs, with traces shown for the signal under 33 µmol m^{-2} s^{-1} (n=5). (K,L) Circadian rhythms of delayed chlorophyll fluorescence at the marimo surface and interior (n=3). (I–L) Red circles represent data and blue lines represent FFT-NLLS waveform fit to data, and grey boxes on x-axes indicate subjective dark period. Asterisks indicate statistically significant differences ($p<0.05$) using (G) Student's t-test and (H) one-way ANOVA with Student-Newman-Keuls post-hoc analysis. Data are mean±s.e.m.

in marimo is also under circadian control. We tested whether the light-induced acquisition of buoyancy by marimo (Figure 1A,B) is circadian regulated. Marimo were entrained to 12 h light/12 h dark cycles and then transferred to continuous dim red light to maintain circadian rhythmicity. When these marimo were exposed to buoyancy-inducing light intensities (20 μmol m^{-2} s^{-1} photon flux density) the marimo became buoyant more rapidly at the start of the subjective day than halfway through the subjective day, during both the second and third days of constant dim light (Figure 1G). This indicates that there is circadian regulation of light-induced buoyancy in marimo. Under light/dark cycles, marimo position within the water column oscillated (Figure S1F in Supplemental Information, published with this article online) whereas marimo remained buoyant under constant light conditions (Figure S1G). Persistent marimo buoyancy under constant light is expected because photosynthesis always occurs under constant light, even when circadian oscillations are present [4].

We identified day–night fluctuations in photosynthetic rates at the marimo surface. Under light/dark cycles, we measured photosynthesis at the marimo surface and interior using a chlorophyll fluorescence instrument with a narrow fibre optic probe. The quantum yield of photosystem II (Y(II)) and electron transport rate (ETR) was significantly greater at the marimo surface during the light period than dark period, under light-limited and light-saturated conditions (Figure 1H). This day–night fluctuation was absent from the marimo interior (Figure 1H). Interior day-time and night-time Y(II) and ETR was comparable to the marimo exterior during the dark period (Figure 1H). Therefore, there are anatomically dependent day–night fluctuations in the photosynthetic rate of marimo.

Next, we identified low amplitude circadian rhythms of photosynthesis in marimo. Marimo were entrained to 12 h light/12 h dark cycles, and intact marimo and dissected marimo interiors then transferred to constant light conditions. Under constant light, rapid photosynthetic light response curves (RLCs) were obtained every 2 h using chlorophyll fluorescence imaging (Figure S1H). To our knowledge, this is the first time the RLCs have been used to study circadian rhythms in plants. Circadian oscillations in Y(II) (essentially the proportion of light entering photochemistry) and non-photochemical quenching (NPQ, thermal dissipation of energy) occurred within a subset of areas of interest analyzed, at many light intensities within the RLCs (Table S1A; Figure 1I,J). More areas of interest were rhythmic at the marimo surface under lower light intensities, whereas this trend was absent from the marimo interior (Table S1). For example, under 33 μmol m^{-2} s^{-1} light the marimo surface was clearly rhythmic whereas the interior was arrhythmic (Figures 1I,J and S1I,J). Circadian period estimates obtained using fast Fourier transform-non linear least squares method (FFT-NLLS) ranged from 20–28 h, depending on light intensity and marimo anatomy (Figures 1I,J and S1I,J). The relative amplitude error (RAE) quantifies the rhythmic robustness of the oscillation, with low RAE indicating a better FFT-NLLS fit to data. Across the light intensities within the RLC timecourses, RAE was in the range of 0.6–0.97 for marimo exterior and 0.45–0.98 for marimo interior (Table S1). This exceeds the RAE of timecourses of promoter–luciferase reporters and chlorophyll fluorescence analysis of *Arabidopsis* (RAE=0.3–0.35 [5]), and is somewhat greater than RAE from chlorophyll fluorescence analysis of barley (RAE=0.2–0.6; some barley varieties 0.4–0.9 [6]). The RAE from marimo might indicate lower amplitude circadian oscillations, and potentially explains the variable phase estimates (Table S1). The mathematically distinct maximum entropy spectral analysis (MESA) also identified circadian rhythms in these photosynthetic parameters (Table S1). The slope of the straight-line portion of the RLC did not oscillate significantly under constant light (Figure S1K). Circadian rhythms were also detected using the delayed chlorophyll fluorescence (DF) method [7] (Figure 1K,L). DF oscillations likely derive from rhythms of charge recombination during photosynthetic electron transport [7]. Together, this identifies low amplitude circadian rhythms of PSII activity in marimo.

Marimo buoyancy might maximise photosynthetic light capture because light penetration decreases with depth. Buoyancy alterations might also rearrange balls within marimo accumulations and underlie seasonal changes in marimo position [3]. There are daily changes in the water column position of jellyfish harbouring endosymbiotic algae [8], and daily valve opening by giant clams exposes endosymbiotic algae to sunlight [9]. Similarly, higher plant organs undergo diel and circadian changes in position [10]. Therefore, circadian and diel movements might optimize photosynthesis in several kingdoms of aquatic and terrestrial organisms.

SUPPLEMENTAL INFORMATION

Supplemental Information contains one figure, one table, and experimental procedures, and can be found with this article online at https://doi.org/10.1016/j.cub.2018.07.027.

REFERENCES

1. Boedeker, C., Eggert, A., Immers, A., and Smets, E. (2010). Global decline of and threats to *Aegagropila linnaei*, with special reference to the lake ball habit. BioScience *60*, 187–198.
2. Togashi, T., Sasaki, H., and Yoshimura, J. (2014). A geometrical approach explains Lake Ball (Marimo) formations in the green alga, *Aegagropila linnaei*. Sci. Rep. *4*, 3761.
3. Acton, E. (1916). On the structure and origin of "Cladophora balls". New Phytol. *15*, 1–10.
4. Dodd, A.N., Salathia, N., Hall, A., Kévei, E., Tóth, R., Nagy, F., Hibberd, J.M., Millar, A.J., and Webb, A.A.R. (2005). Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. Science *309*, 630.
5. Litthauer, S., Battle Martin, W., Lawson, T., and Jones Matthew, A. (2015). Phototropins maintain robust circadian oscillation of PSII operating efficiency under blue light. Plant J. *83*, 1034–1045.
6. Dakhiya, Y., Hussien, D., Fridman, E., Kiflawi, M., and Green, R. (2017). Correlations between circadian rhythms and growth in challenging environments. Plant Physiol. *173*, 1724.
7. Gould, P.D., Diaz, P., Hogben, C., Kusakina, J., Salem, R., Hartwell, J., and Hall, A. (2009). Delayed fluorescence as a universal tool for the measurement of circadian rhythms in higher plants. Plant J. *58*, 893–901.
8. Hamner, W.M. (1997). Sensory ecology of scyphmedusae. In Zooplankton: Sensory Ecology and Physiology, P.H. Lenz, D.K. Hartline, J.E. Purcell and D.L. Macmillan, eds. (Gordon and Breach).
9. Morton, B. (2009). The diurnal rhythm and the processes of feeding and digestion in *Tridacna crocea* (BivaMa: Tridacnidae). J. Zool. *185*, 371–387.
10. Darwin, C.R. (1880). The Power of Movement in Plants (John Murray).

School of Biological Sciences, Life Sciences Building, University of Bristol, Bristol BS8 1TQ, UK.
[1]Lead contact.
*E-mail: antony.dodd@bristol.ac.uk

Received: 6 May 2018 | Revised: 16 May 2018 | Accepted: 16 May 2018

DOI: 10.1111/pce.13346

COMMENTARY

WILEY Plant, Cell & Environment

New connections between circadian rhythms, photosynthesis, and environmental adaptation

Dora L. Cano-Ramirez | Antony N. Dodd

School of Biological Sciences, Life Sciences Building, University of Bristol, 24 Tyndall Avenue, Bristol BS8 1TQ, UK

Correspondence
Antony N. Dodd, School of Biological Sciences, Life Sciences Building, University of Bristol, 24 Tyndall Avenue, Bristol BS8 1TQ, UK.
Email: antony.dodd@bristol.ac.uk

1 | MAIN TEXT

Daily cycles of day and night create environmental changes that challenge plants in a variety of ways. For example, there are daily cycles in the availability of light for photosynthesis, cycles of temperature and humidity, and daily changes in herbivore activity. Circadian rhythms appear to have evolved as an adaptation to predictable daily changes in the environment and have incredibly pervasive effects upon plant physiology and development. There is circadian regulation of photosynthesis, stomatal opening, flower position, scent emission, metabolism, growth, and defence (Hennessey & Field, 1991; Millar, 2016). Furthermore, circadian rhythms confer a fitness advantage to plants (Dodd et al., 2005; Green, Tingay, Wang, & Tobin, 2002). Circadian rhythms are generated by the cellular circadian oscillator, which in higher plants comprises a series of interlocked transcription-translation feedback loops that complete one cycle of transcription and translation in about 24 hr. Mechanisms have also evolved that adjust the circadian phase to match the environmental phase (known as entrainment), and that couple the circadian oscillator with circadian-regulated metabolism and physiology.

The mechanisms that underlie the circadian regulation of photosynthesis have remained somewhat mysterious. The nuclear-encoded circadian oscillator regulates light harvesting, CO_2 uptake, and stomatal opening because mutations to the circadian oscillator that change its free-running circadian period lead to equivalent changes in the period of these physiological processes (Dodd et al., 2005; Dodd, Parkinson, & Webb Alex, 2004; Litthauer, Battle Martin, Lawson, & Jones Matthew, 2015; Somers, Webb, Pearson, & Kay, 1998). This might be due to circadian rhythms of the accumulation of many transcripts encoding proteins associated with the light-dependent and light-independent reactions of photosynthesis (Harmer et al., 2000). There are also circadian rhythms of chloroplast gene transcription (Noordally et al., 2013), but the contribution of these rhythms to the circadian timing of PSII activity is unclear. Rhythms of phosphorylation of photosynthetic proteins (Choudhary, Nomura, Wang, Nakagami, & Somers, 2015) might also contribute to the circadian regulation of photosynthetic physiology. These rhythms of photosynthesis can be investigated using chlorophyll fluorescence techniques. There are circadian rhythms in the operating efficiency of PSII (Fv'/Fm') which, intriguingly, requires the phototropin blue light photoreceptors (Litthauer et al., 2015). There are also circadian rhythms of non-photochemical quenching (NPQ; Dakhiya, Hussien, Fridman, Kiflawi, & Green, 2017) and delayed chlorophyll fluorescence (Gould et al., 2009). Remarkably, circadian rhythms of oxygen evolution persist in the alga *Acetabularia* following removal of the nucleus (Sweeney & Haxo, 1961), raising the tantalizing possibility of chloroplast-autonomous circadian rhythms. It seems that a range of rhythmic processes might act in combination to elicit the syndrome of the circadian regulation of photosynthesis, with techniques such as modulated chlorophyll fluorescence being attractive to dissect facets of this process.

In this issue of *Plant, Cell, and Environment*, Yarkhunova et al. (2018) investigated the influence of circadian regulation upon various aspects of photosynthetic light harvesting. Using *Arabidopsis thaliana*, they found that mutations that alter the circadian period decrease the operating efficiency of Photosystem II and increase the non-photochemical quenching (NPQ) of chlorophyll fluorescence. The authors also took a recombinant inbred line (RIL) population of Arabidopsis, produced from a cross between the Wassilewskija and Landsberg *erecta* accessions. Using this population, Yarkhunova et al. (2018) identified that circadian periods closer to 24 hr were associated with greater operating efficiency of PSII, whereas longer periods within the population associated with decreased PSII operating efficiency. Conversely, NPQ was positively correlated with circadian periods in the range 24.5–27 hr. Overall, this suggests that when the circadian period is longer than 24 hr, there is a reduction in the efficiency of photosynthetic light harvesting and this is accompanied by an increase in thermal dissipation of energy. Interestingly, the authors identified that a particular component of thermal dissipation, state-transition-related quenching (*qT*), was increased when the circadian period was closer to 24 hr. This suggests

a role for circadian regulation in the regulation of specific aspects of non-photochemical quenching, which is supported by circadian regulation of some transcripts associated with *qT*.

One potential explanation for decreased PSII operating efficiency and increased NPQ in plants having a circadian period divergent from 24 hr relates to the concept of "circadian resonance" (Figure 1). Circadian resonance is a concept that refers to the phase relationship between circadian-regulated cellular processes and the environmental day–night cycles (Ouyang, Andersson, Kondo, Golden, & Johnson, 1998; Pittendrigh & Minis, 1972). Within this concept, when the cellular circadian period is similar to the environmental period, the circadian-regulated processes will have phases that are appropriate for the day–night cycle. Under these circumstances, the endogenous circadian oscillation is resonant with the environmental rhythm (Figure 1). In comparison, if the circadian period deviates from the environmental period, the phase of cellular processes will become suboptimally timed relative to the environment and lose the benefits conferred by circadian regulation (Figure 1). In the context of the regulation of PSII, circadian resonance might ensure that proteins associated with the regulation of NPQ or proteins that are turned over rapidly in the light, such as PSII D1 and PSII D2, accumulate at those times of day with greatest light intensity. Alternative explanations for the reduction in PSII operating efficiency and increase in NPQ when the circadian period is mismatched from the environmental period might include effects of circadian period upon stomatal function, leaf internal structure, subcellular positioning of chloroplasts, and changes in the concentration of regulatory carbohydrates.

The study of Yarkhunova et al. (2018) has interesting implications. For example, the genetic diversity generated during crop breeding will cause variation in circadian period that could alter photosynthetic performance (Yarkhunova et al., 2016). Whether this affects overall crop performance might be context dependent; for example, the selection for a longer circadian period during tomato domestication is thought to be advantageous (Müller et al., 2015), so it would be interesting to know whether this longer circadian period reduces aspects of PSII operating efficiency despite providing an adaptation to the longer

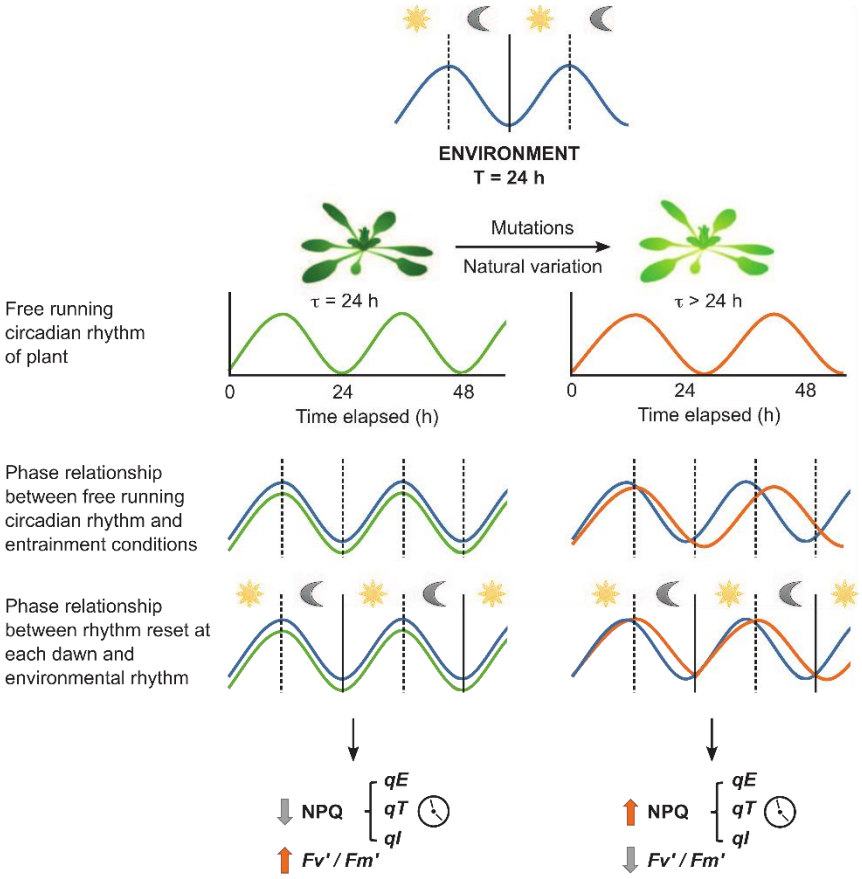


FIGURE 1 Potential consequence of circadian period (τ) that deviates from the environmental period (T). When T and τ are similar the phase of the free running circadian rhythm remains phase-matched with the previous entrainment conditions, and under light/dark cycles the circadian rhythm provides a good estimate of the time of day. When τ deviates from T (in this example, τ > T), under free running conditions the phase of the free running rhythm will deviate increasingly from the phase of entrainment. This means that when τ > T under light/dark cycles, the phase of rhythmic processes will deviate increasingly from the environmental cycle until it is reset by the next zeitgeber (e.g., dawn) (Dodd, Dalchau, Gardner, Baek, & Webb, 2013). It is under conditions where τ deviates from T, in both circadian oscillator mutants and through genetic diversity, that the operating efficiency of PSII is reduced and NPQ is increased (Yarkhunova et al., 2016). Blue line represents phase of environmental day/night cycle where *T* = 24 hr, and green and orange lines represent endogenous rhythms within plants where τ = 24 hr and τ > 24 hr, respectively. *Fv′/Fm′* = maximum efficiency of PSII in light; NPQ = non photochemical quenching of chlorophyll fluorescence; *qE* = pH-dependent quenching; *qT* = state-transition quenching; *qI* = quenching related to photo inhibition [Colour figure can be viewed at wileyonlinelibrary.com]

photoperiods of northern-latitude summers. Given that there is also natural variation in circadian period within isogenic populations (Swarup et al., 2002), there might be continuous genetic and epigenetic selection within nature for a certain range of circadian periods, due to the deleterious effects upon plant performance of a circadian period that is divergent from the environmental period. Indeed, identification of individual seedlings having a particular range of circadian periods, using chlorophyll fluorescence analysis in lettuce, has provided a tool to enhance yield consistency within vertical agriculture systems (Moriyuki & Fukuda, 2016). Together, this suggests that deliberate selection of defined circadian traits will form an important part of future crop breeding.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

ORCID

Antony N. Dodd http://orcid.org/0000-0001-6859-0105

REFERENCES

Choudhary, M. K., Nomura, Y., Wang, L., Nakagami, H., & Somers, D. E. (2015). Quantitative circadian phosphoproteomic analysis of Arabidopsis reveals extensive clock control of key components in physiological, metabolic, and signaling pathways. *Molecular & Cellular Proteomics*, *14*(8), 2243–2260.

Dakhiya, Y., Hussien, D., Fridman, E., Kiflawi, M., & Green, R. (2017). Correlations between circadian rhythms and growth in challenging environments. *Plant Physiology*, *173*(3), 1724–1734.

Dodd, A. N., Dalchau, N., Gardner, M. J., Baek, S. J., & Webb, A. A. (2013). The circadian clock has transient plasticity of period and is required for timing of nocturnal processes in Arabidopsis. *New Phytologist*, *201*(1), 168–179.

Dodd, A. N., Parkinson, K., & Webb Alex, A. R. (2004). Independent circadian regulation of assimilation and stomatal conductance in the *ztl*-1 mutant of Arabidopsis. *New Phytologist*, *162*(1), 63–70.

Dodd, A. N., Salathia, N., Hall, A., Kévei, E., Tóth, R., Nagy, F., … Webb, A. A. R. (2005). Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science*, *309*(5734), 630–633.

Gould, P. D., Diaz, P., Hogben, C., Kusakina, J., Salem, R., Hartwell, J., & Hall, A. (2009). Delayed fluorescence as a universal tool for the measurement of circadian rhythms in higher plants. *The Plant Journal*, *58*(5), 893–901.

Green, R. M., Tingay, S., Wang, Z.-Y., & Tobin, E. M. (2002). Circadian rhythms confer a higher level of fitness to Arabidopsis plants. *Plant Physiology*, *129*(2), 576–584.

Harmer, S. L., Hogenesch, J. B., Straume, M., Chang, H.-S., Han, B., Zhu, T., … Kay, S. A. (2000). Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science*, *290*(5499), 2110–2113.

Hennessey, T. L., & Field, C. B. (1991). Circadian rhythms in photosynthesis: Oscillations in carbon assimilation and stomatal conductance under constant conditions. *Plant Physiology*, *96*(3), 831–836.

Litthauer, S., Battle Martin, W., Lawson, T., & Jones Matthew, A. (2015). Phototropins maintain robust circadian oscillation of PSII operating efficiency under blue light. *The Plant Journal*, *83*(6), 1034–1045.

Millar, A. J. (2016). The intracellular dynamics of circadian clocks reach for the light of ecology and evolution. *Annual Review of Plant Biology*, *67*(1), 595–618.

Moriyuki, S., & Fukuda, H. (2016). High-throughput growth prediction for *Lactuca sativa* L. seedlings using chlorophyll fluorescence in a plant factory with artificial lighting. *Frontiers in Plant Science*, *7*, 394.

Müller, N. A., Wijnen, C. L., Srinivasan, A., Ryngajllo, M., Ofner, I., Lin, T., … Jiménez-Gómez, J. M. (2015). Domestication selected for deceleration of the circadian clock in cultivated tomato. *Nature Genetics*, *48*, 89–93.

Noordally, Z. B., Ishii, K., Atkins, K. A., Wetherill, S. J., Kusakina, J., Walton, E. J., … Dodd, A. N. (2013). Circadian control of chloroplast transcription by a nuclear-encoded timing signal. *Science*, *339*(6125), 1316–1319.

Ouyang, Y., Andersson, C. R., Kondo, T., Golden, S. S., & Johnson, C. H. (1998). Resonating circadian clocks enhance fitness in cyanobacteria. *Proceedings of the National Academy of Sciences*, *95*(15), 8660–8664.

Pittendrigh, C. S., & Minis, D. H. (1972). Circadian systems: Longevity as a function of circadian resonance in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*, *69*, 1537–1539.

Somers, D. E., Webb, A. A., Pearson, M., & Kay, S. A. (1998). The short-period mutant, *toc1*-1, alters circadian clock regulation of multiple outputs throughout development in *Arabidopsis thaliana*. *Development*, *125*(3), 485–494.

Swarup, K., Alonso-Blanco, C., Lynn James, R., Michaels Scott, D., Amasino Richard, M., Koornneef, M., & Millar Andrew, J. (2002). Natural allelic variation identifies new genes in the Arabidopsis circadian system. *The Plant Journal*, *20*(1), 67–77.

Sweeney, B. M., & Haxo, F. T. (1961). Persistence of a photosynthetic rhythm in enucleated Acetabularia. *Science*, *134*(3487), 1361–1363.

Yarkhunova, Y., Guadagno, C. R., Rubin, M., Ewers, B. E., Davis, S. J., Weinig, C. (2018). Circadian rhythms are associated with variation in photosystem II function and photoprotective mechanisms. Plant, Cell & Environment.

Yarkhunova, Y., Edwards Christine, E., Ewers Brent, E., Baker Robert, L., Aston Timothy, L., McClung, C. R., … Weinig, C. (2016). Selection during crop diversification involves correlated evolution of the circadian clock and ecophysiological traits in *Brassica rapa*. *New Phytologist*, *210*(1), 133–144.

How to cite this article: Cano-Ramirez DL, Dodd AN. New connections between circadian rhythms, photosynthesis, and environmental adaptation. *Plant Cell Environ*. 2018;1–3. https://doi.org/10.1111/pce.13346

9 REFERENCES

ACTON, E. 1916. On the structure and origin of “Cladophora balls”. *New Phytologist,* 15**,** 97-103.

AIKAWA, S., KOBAYASHI, M. J., SATAKE, A., SHIMIZU, K. K. & KUDOH, H. 2010. Robust control of the seasonal expression of the Arabidopsis FLC gene in a fluctuating environment. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 11632-11637.

ALLEN, J. F. 2015. Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** 10231-10238.

ANGEL, A., SONG, J., DEAN, C. & HOWARD, M. 2011. A Polycomb-based switch underlying quantitative epigenetic memory. *Nature,* 476**,** 105.

ARNOSTI, D. N. & CHAMBERLIN, M. J. 1989. Secondary sigma factor controls transcription of flagellar and chemotaxis genes in Escherichia coli. *Proceedings of the National Academy of Sciences,* 86**,** 830-834.

ARTUS, N. N., UEMURA, M., STEPONKUS, P. L., GILMOUR, S. J., LIN, C. & THOMASHOW, M. F. 1996. Constitutive expression of the cold-regulated *Arabidopsis thaliana* COR15a gene affects both chloroplast and protoplast freezing tolerance. *Proceedings of the National Academy of Sciences,* 93**,** 13404.

ATKINS, K. A. & DODD, A. N. 2014. Circadian regulation of chloroplasts. *Current Opinion in Plant Biology,* 21**,** 43-50.

BARAJAS-LÓPEZ, J. D. D., BLANCO, N. E. & STRAND, Å. 2013. Plastid-to-nucleus communication, signals controlling the running of the plant cell. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1833**,** 425-437.

BARNES, A. C., BENNING, C. & ROSTON, R. L. 2016a. Chloroplast membrane remodeling during freezing stress is accompanied by cytoplasmic acidification activating SENSITIVE TO FREEZING2. *Plant Physiology,* 171**,** 2140.

BARNES, A. C., BENNING, C. & ROSTON, R. L. 2016b. Chloroplast membrane remodeling during freezing stress is accompanied by cytoplasmic acidification activating SENSITIVE TO FREEZING2. *Plant Physiology,* 171**,** 2140-9.

BASTOW, R., MYLNE, J. S., LISTER, C., LIPPMAN, Z., MARTIENSSEN, R. A. & DEAN, C. 2004. Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature,* 427**,** 164.

BELBIN, F., NOORDALLY ZEENAT, B., WETHERILL SARAH, J., ATKINS KELLY, A., FRANKLIN KEARA, A. & DODD ANTONY, N. 2016. Integration of light and circadian signals that regulate chloroplast transcription by a nuclear-encoded sigma factor. *New Phytologist,* 213**,** 727-738.

BELL-PEDERSEN, D., CASSONE, V. M., EARNEST, D. J., GOLDEN, S. S., HARDIN, P. E., THOMAS, T. L. & ZORAN, M. J. 2005. Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature Reviews Genetics,* 6**,** 544-56.

BENNING, C. 2009. Mechanisms of lipid transport involved in organelle biogenesis in plant cells. *Annual Review of Cell and Developmental Biology,* 25**,** 71-91.

BIENIAWSKA, Z., ESPINOZA, C., SCHLERETH, A., SULPICE, R., HINCHA, D. K. & HANNAH, M. A. 2008. Disruption of the Arabidopsis circadian clock is responsible for extensive variation in the cold-responsive transcriptome. *Plant Physiology,* 147**,** 263-79.

BJORNSON, M., BALCKE, G. U., XIAO, Y., SOUZA, A., WANG, J.-Z., ZHABINSKAYA, D., TAGKOPOULOS, I., TISSIER, A. & DEHESH, K. 2017. Integrated omics analyses of retrograde signaling mutant delineate interrelated stress-response strata. *Plant Journal,* 91**,** 70-84.

BLANE, A. & FANUCCHI, S. 2015. Effect of pH on the structure and DNA binding of the FOXP2 forkhead domain. *Biochemistry,* 54**,** 4001-4007.

BODE, R., IVANOV, A. G. & HÜNER, N. P. A. 2016. Global transcriptome analyses provide evidence that chloroplast redox state contributes to intracellular as well as long-distance signalling in response to stress and acclimation in Arabidopsis. *Photosynthesis Research,* 128**,** 287-312.

BOEDEKER, C. (2010) *Phylogenetic, taxonomic and biogeographical studies in the Pithophoraceae (Cladophorales, Chlorophyta)*, Netherlands Centre for Biodiversity Naturalis (section National Herbarium of the Netherlands), Leiden University Branch.

BOEDEKER, C., EGGERT, A., IMMERS, A. & SMETS, E. 2010. Global decline of and threats to *Aegagropila linnaei*, with special reference to the lake ball habit. *BioScience,* 60**,** 187-198.

BORKHSENIOUS, O. N., MASON, C. B. & MORONEY, J. V. 1998. The intracellular localization of ribulose-1,5-bisphosphate Carboxylase/Oxygenase in *Chlamydomonas reinhardtii*. *Plant Physiology,* 116**,** 1585.

BRADBEER, J. W., ATKINSON, Y. E., BÖRNER, T. & HAGEMANN, R. 1979. Cytoplasmic synthesis of plastid polypeptides may be controlled by plastid-synthesised RNA. *Nature,* 279**,** 816.

BROWN, B. A., CLOIX, C., JIANG, G. H., KAISERLI, E., HERZYK, P., KLIEBENSTEIN, D. J. & JENKINS, G. I. 2005. A UV-B-specific signaling component orchestrates

plant UV protection. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 18225-30.

BROWN, B. A. & JENKINS, G. I. 2008. UV-B signaling pathways with different fluence-rate response profiles are distinguished in mature Arabidopsis leaf tissue by requirement for UVR8, HY5, and HYH. *Plant Physiology,* 146**,** 576-88.

BRUNKARD, JACOB O. & BURCH-SMITH, TESSA M. 2018. Ties that bind: the integration of plastid signalling pathways in plant cell metabolism. *Essays In Biochemistry,* 62**,** 95.

BRÄUTIGAM, K., DIETZEL, L., KLEINE, T., STRÖHER, E., WORMUTH, D., DIETZ, K.-J., RADKE, D., WIRTZ, M., HELL, R., DÖRMANN, P., NUNES-NESI, A., SCHAUER, N., FERNIE, A. R., OLIVER, S. N., GEIGENBERGER, P., LEISTER, D. & PFANNSCHMIDT, T. 2009. Dynamic plastid redox signals integrate gene expression and metabolism to induce distinct metabolic states in photosynthetic acclimation in Arabidopsis. *Plant Cell,* 21**,** 2715.

BURNHAM, K. P. & ANDERSON, D. R. 2004. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research,* 33**,** 261-304.

BYRNE, T. E., WELLS, M. R. & JOHNSON, C. H. 1992. Circadian rhythms of chemotaxis to ammonium and of methylammonium uptake in Chlamydomonas. *Plant Physiology,* 98**,** 879.

BÉDARD, J. & JARVIS, P. 2005. Recognition and envelope translocation of chloroplast preproteins. *Journal of Experimental Botany,* 56**,** 2287-2320.

CALIXTO, C. P. G., GUO, W., JAMES, A. B., TZIOUTZIOU, N. A., ENTIZNE, J. C., PANTER, P. E., KNIGHT, H., NIMMO, H., ZHANG, R. & BROWN, J. W. S. 2018. Rapid and dynamic alternative splicing impacts the Arabidopsis cold response transcriptome. *Plant Cell*.

CANO-RAMIREZ, D. L., SASKIA DE FRAINE, T., GRIFFITHS, O. G. & DODD, A. N. 2018. Photosynthesis and circadian rhythms regulate the buoyancy of marimo lake balls. *Current Biology,* 28**,** R869-R870.

CATALA, R., MEDINA, J. & SALINAS, J. 2011. Integration of low temperature and light signaling during cold acclimation response in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 16475-80.

CHEHAB, E. W., RAMAN, G., WALLEY, J. W., PEREA, J. V., BANU, G., THEG, S. & DEHESH, K. 2006. Rice HYDROPEROXIDE LYASES with unique expression patterns generate distinct aldehyde signatures in Arabidopsis. *Plant Physiology,* 141**,** 121-134.

CHEN, C., XIAO, Y. G., LI, X. & NI, M. 2012. Light-regulated stomatal aperture in Arabidopsis. *Molecular Plant* 5**,** 566-72.

CHEN, H., ZHANG, B., HICKS, L. M. & XIONG, L. 2011. A nucleotide metabolite controls stress-responsive gene expression and plant development. *PLOS One,* 6**,** e26661.

CHOUDHARY, M. K., NOMURA, Y., WANG, L., NAKAGAMI, H. & SOMERS, D. E. 2015. Quantitative Circadian Phosphoproteomic Analysis of Arabidopsis Reveals Extensive Clock Control of Key Components in Physiological, Metabolic, and Signaling Pathways. *Molecular & Cellular Proteomics,* 14**,** 2243.

CHOW, B. Y., SANCHEZ, S. E., BRETON, G., PRUNEDA-PAZ, J. L., KROGAN, N. T. & KAY, S. A. 2014. Transcriptional regulation of LUX by CBF1 mediates cold input to the circadian clock in Arabidopsis. *Current Biology,* 24**,** 1518-24.

COVINGTON, M. F., PANDA, S., LIU, X. L., STRAYER, C. A., WAGNER, D. R. & KAY, S. A. 2001. ELF3 modulates resetting of the circadian clock in Arabidopsis. *Plant Cell,* 13**,** 1305-1316.

DALCHAU, N., BAEK, S. J., BRIGGS, H. M., ROBERTSON, F. C., DODD, A. N., GARDNER, M. J., STANCOMBE, M. A., HAYDON, M. J., STAN, G.-B., GONÇALVES, J. M. & WEBB, A. A. R. 2011. The circadian oscillator gene GIGANTEA mediates a long-term response of the *Arabidopsis thaliana* circadian clock to sucrose. *Proceedings of the National Academy of Sciences,* 108**,** 5104.

DANIELL, H., LIN, C.-S., YU, M. & CHANG, W.-J. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biology,* 17**,** 134.

DE LONGEVIALLE, A. F., SMALL, I. D. & LURIN, C. 2010. Nuclearly encoded splicing factors implicated in RNA splicing in higher plant organelles. *Molecular Plant,* 3**,** 691-705.

DELWICHE, CHARLES F. & COOPER, ENDYMION D. 2015. The evolutionary origin of a terrestrial flora. *Current Biology,* 25**,** R899-R910.

DODD, A. N., BELBIN, F. E., FRANK, A. & WEBB, A. A. 2015. Interactions between circadian clocks and photosynthesis for the temporal and spatial coordination of metabolism. *Frontiers in Plant Science,* 6**,** 245.

DODD, A. N., SALATHIA, N., HALL, A., KEVEI, E., TOTH, R., NAGY, F., HIBBERD, J. M., MILLAR, A. J. & WEBB, A. A. 2005. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science,* 309**,** 630-3.

DOLINSKY, T. J., NIELSEN, J. E., MCCAMMON, J. A. & BAKER, N. A. 2004. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Research,* 32**,** W665-W667.

DONG, M. A., FARRE, E. M. & THOMASHOW, M. F. 2011. Circadian clock-associated 1 and late elongated hypocotyl regulate expression of the C-repeat binding factor (CBF) pathway in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 7241-6.

DORRELL, R. G., DREW, J., NISBET, R. E. R. & HOWE, C. J. 2014. Evolution of Chloroplast Transcript Processing in Plasmodium and Its Chromerid Algal Relatives. *PLOS Genetics,* 10**,** e1004008.

DORRELL, R. G. & HOWE, C. J. 2012. What makes a chloroplast? Reconstructing the establishment of photosynthetic symbioses. *Journal of Cell Science,* 125**,** 1865-1875.

DUREK, P., SCHMIDT, R., HEAZLEWOOD, J. L., JONES, A., MACLEAN, D., NAGEL, A., KERSTEN, B. & SCHULZE, W. X. 2010. PhosPhAt: the *Arabidopsis thaliana* phosphorylation site database. An update. *Nucleic Acids Research,* 38**,** D828-D834.

ECKARDT, N. A. 2006. A wheel within a wheel: Temperature compensation of the circadian clock. *Plant Cell,* 18**,** 1105.

EDGAR, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research,* 32**,** 1792-1797.

EDGAR, R. S., GREEN, E. W., ZHAO, Y., VAN OOIJEN, G., OLMEDO, M., QIN, X., XU, Y., PAN, M., VALEKUNJA, U. K., FEENEY, K. A., MAYWOOD, E. S., HASTINGS, M. H., BALIGA, N. S., MERROW, M., MILLAR, A. J., JOHNSON, C. H., KYRIACOU, C. P., O'NEILL, J. S. & REDDY, A. B. 2012. Peroxiredoxins are conserved markers of circadian rhythms. *Nature,* 485**,** 459-64.

ENDO, M., SHIMIZU, H., NOHALES, M. A., ARAKI, T. & KAY, S. A. 2014. Tissue-specific clocks in Arabidopsis show asymmetric coupling. *Nature,* 515**,** 419-422.

ERIKSSON, M. E. & WEBB, A. A. 2011. Plant cell responses to cold are all about timing. *Current Opinion in Plant Biology,* 14**,** 731-7.

ESPINAS, N. A., KOBAYASHI, K., SATO, Y., MOCHIZUKI, N., TAKAHASHI, K., TANAKA, R. & MASUDA, T. 2016. Allocation of heme is differentially regulated by ferrochelatase isoforms in Arabidopsis cells. *Frontiers in Plant Science,* 7.

ESTAVILLO, G. M., CRISP, P. A., PORNSIRIWONG, W., WIRTZ, M., COLLINGE, D., CARRIE, C., GIRAUD, E., WHELAN, J., DAVID, P., JAVOT, H., BREARLEY, C., HELL, R., MARIN, E. & POGSON, B. J. 2011. Evidence for a SAL1-PAP chloroplast retrograde pathway that functions in drought and high light signaling in Arabidopsis. *Plant Cell,* 23**,** 3992.

FARINAS, B. & MAS, P. 2011. Functional implication of the MYB transcription factor RVE8/LCL5 in the circadian control of histone acetylation. *Plant Journal,* 66**,** 318-329.

FAST, N. M., KISSINGER, J. C., ROOS, D. S. & KEELING, P. J. 2001. Nuclear-encoded, plastid-targeted genes suggest a single common origin for Apicomplexan and Dinoflagellate plastids. *Molecular Biology and Evolution,* 18**,** 418-426.

FAVORY, J. J., KOBAYSHI, M., TANAKA, K., PELTIER, G., KREIS, M., VALAY, J. G. & LERBS-MACHE, S. 2005. Specific function of a plastid sigma factor for ndhF gene transcription. *Nucleic Acids Research,* 33**,** 5991-9.

FEHÉR, B., KOZMA-BOGNÁR, L., KEVEI, É., HAJDU, A., BINKERT, M., DAVIS, S. J., SCHÄFER, E., ULM, R. & NAGY, F. 2011. Functional interaction of the circadian clock and UV RESISTANCE LOCUS8-controlled UV-B signaling pathways in *Arabidopsis thaliana*. *Plant Journal,* 67**,** 37-48.

FINN, R. D., COGGILL, P., EBERHARDT, R. Y., EDDY, S. R., MISTRY, J., MITCHELL, A. L., POTTER, S. C., PUNTA, M., QURESHI, M., SANGRADOR-VEGAS, A., SALAZAR, G. A., TATE, J. & BATEMAN, A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research,* 44**,** D279-D285.

FLORES-PÉREZ, Ú., PÉREZ-GIL, J., CLOSA, M., WRIGHT, L. P., BOTELLA-PAVÍA, P., PHILLIPS, M. A., FERRER, A., GERSHENZON, J. & RODRÍGUEZ-CONCEPCIÓN, M. 2010. PLEIOTROPIC REGULATORY LOCUS 1 (PRL1) Integrates the Regulation of Sugar Responses with Isoprenoid Metabolism in Arabidopsis. *Molecular Plant,* 3**,** 101-112.

FOWLER, S. G., COOK, D. & THOMASHOW, M. F. 2005. Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock. *Plant Physiology,* 137**,** 961-8.

FRANKLIN, K. A. 2008. Shade avoidance. *New Phytologist,* 179**,** 930-44.

FRANKLIN, K. A. & WHITELAM, G. C. 2007. Light-quality regulation of freezing tolerance in *Arabidopsis thaliana*. *Nature Genetics,* 39**,** 1410-1413.

FUJII, Y. & KODAMA, Y. 2018. Refinements to light sources used to analyze the chloroplast cold-avoidance response over the past century. *Plant Signaling & Behavior,* 13**,** e1411452.

FUJII, Y., TANAKA, H., KONNO, N., OGASAWARA, Y., HAMASHIMA, N., TAMURA, S., HASEGAWA, S., HAYASAKI, Y., OKAJIMA, K. & KODAMA, Y. 2017. Phototropin perceives temperature based on the lifetime of its photoactivated state. *Proceedings of the National Academy of Sciences,* 114**,** 9206.

GILMOUR, S. J., ZARKA, D. G., STOCKINGER, E. J., SALAZAR, M. P., HOUGHTON, J. M. & THOMASHOW, M. F. 1998. Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *Plant Journal,* 16**,** 433-442.

GIORDANO, M., BEARDALL, J. & RAVEN, J. A. 2005. CO_2 Concentrating mechanisms in algae: Mechanisms, environmental modulation, and evolution. *Annual Review of Plant Biology,* 56**,** 99-131.

GOULARD, F., LÜNING, K. & JACOBSEN, S. 2004. Circadian rhythm of photosynthesis and concurrent oscillations of transcript abundance of photosynthetic genes in the marine red alga Grateloupia turuturu. *European Journal of Phycology,* 39**,** 431-437.

GOULD, P. D., LOCKE, J. C., LARUE, C., SOUTHERN, M. M., DAVIS, S. J., HANANO, S., MOYLE, R., MILICH, R., PUTTERILL, J., MILLAR, A. J. & HALL, A. 2006. The molecular basis of temperature compensation in the Arabidopsis circadian clock. *Plant Cell,* 18**,** 1177-87.

GRAY, J. A., SHALIT-KANEH, A., CHU, D. N., HSU, P. Y. & HARMER, S. L. 2017. The REVEILLE clock genes inhibit growth of juvenile and adult plants by control of cell size. *Plant Physiology,* 173**,** 2308-2322.

GRAY, J. C., SORNARAJAH, R., ZABRON, A. A., DUCKETT, C. M. & KHAN, M. S. 1995. Chloroplast control of nuclear gene expression. *In:* MATHIS, P. (ed.) *Photosynthesis: From light to biosphere.* Dordrecht, Netherlands: Kluwer.

GRAY, J. C., SULLIVAN, J. A., WANG, J.-H., JEROME, C. A. & MACLEAN, D. 2003. Coordination of plastid and nuclear gene expression. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 358**,** 135-145.

GREEN, R. M., TINGAY, S., WANG, Z. Y. & TOBIN, E. M. 2002. Circadian rhythms confer a higher level of fitness to Arabidopsis plants. *Plant Physiology,* 129**,** 576-584.

GUTKNECHT, J., BISSON, M. A. & TOSTESON, F. C. 1977. Diffusion of carbon dioxide through lipid bilayer membranes. Effects of carbonic anhydrase, bicarbonate, and unstirred layers. *The Journal of General Physiology,* 69**,** 779-794.

GY, I., GASCIOLLI, V., LAURESSERGUES, D., MOREL, J.-B., GOMBERT, J., PROUX, F., PROUX, C., VAUCHERET, H. & MALLORY, A. C. 2007. Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. *Plant Cell,* 19**,** 3451.

HALFORD, S. E. & MARKO, J. F. 2004. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Research,* 32**,** 3040-3052.

HANAOKA, M., KANAMARU, K., FUJIWARA, M., TAKAHASHI, H. & TANAKA, K. 2005. Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. *EMBO Reports,* 6**,** 545-550.

HANAOKA, M., KANAMARU, K., TAKAHASHI, H. & TANAKA, K. 2003. Molecular genetic analysis of chloroplast gene promoters dependent on SIG2, a nucleus-encoded sigma factor for the plastid-encoded RNA polymerase, in *Arabidopsis thaliana*. *Nucleic Acids Research,* 31**,** 7090-7098.

HANAOKA, M., KATO, M., ANMA, M. & TANAKA, K. 2012. SIG1, a sigma factor for the chloroplast RNA polymerase, differently associates with multiple DNA regions in the chloroplast chromosomes in vivo. *International Journal of Molecular Sciences,* 13**,** 12182-94.

HARMER, S. L., HOGENESCH, J. B., STRAUME, M., CHANG, H. S., HAN, B., ZHU, T., WANG, X., KREPS, J. A. & KAY, S. A. 2000. Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science,* 290**,** 2110-2113.

HARPENSLAGER, S. F., SMOLDERS, A. J. P., KIESKAMP, A. A. M., ROELOFS, J. G. M. & LAMERS, L. P. M. 2015. To float or not to float: How interactions between light and dissolved inorganic carbon species determine the buoyancy of *Stratiotes aloides*. *PLOS One,* 10**,** e0124026.

HAVAUX, M. 2013. Carotenoid oxidation products as stress signals in plants. *Plant Journal,* 79**,** 597-606.

HELMANN, J. D. 2002. The extracytoplasmic function (ECF) sigma factors. *Advances in microbial physiology* 46**,** 47 -110.

HEMSLEY, P. A., HURST, C. H., KALIYADASA, E., LAMB, R., KNIGHT, M. R., DE COTHI, E. A., STEELE, J. F. & KNIGHT, H. 2014. The Arabidopsis mediator complex subunits MED16, MED14, and MED2 regulate mediator and RNA polymerase II recruitment to CBF-responsive cold-regulated genes. *Plant Cell,* 26 465-484.

HEPWORTH, J., ANTONIOU-KOUROUNIOTI, R. L., BLOOMER, R. H., SELGA, C., BERGGREN, K., COX, D., COLLIER HARRIS, B. R., IRWIN, J. A., HOLM, S., SÄLL, T., HOWARD, M. & DEAN, C. 2018. Absence of warmth permits epigenetic memory of winter in Arabidopsis. *Nature Communications,* 9**,** 639.

HIDEMA, J., KUMAGAI, T., SUTHERLAND, J. C. & SUTHERLAND, B. M. 1997. Ultraviolet B-sensitive rice cultivar deficient in cyclobutyl pyrimidine dimer repair. *Plant Physiology,* 113**,** 39.

HIDEMA, J. U. N. & KUMAGAI, T. 2006. Sensitivity of rice to Ultraviolet-B radiation. *Annals of Botany,* 97**,** 933-942.

HOTTA, C. T., GARDNER, M. J., HUBBARD, K. E., BAEK, S. J., DALCHAU, N., SUHITA, D., DODD, A. N. & WEBB, A. A. 2007. Modulation of environmental responses of plants by circadian clocks. *Plant, Cell & Environment,* 30**,** 333-49.

HSU, P. Y., DEVISETTY, U. K. & HARMER, S. L. 2013. Accurate timekeeping is controlled by a cycling activator in Arabidopsis. *eLife,* 2**,** e00473.

HUERTA-CEPAS, J., CAPELLA-GUTIÉRREZ, S., PRYSZCZ, L. P., MARCET-HOUBEN, M. & GABALDÓN, T. 2014. PhylomeDB v4: zooming into the plurality of evolutionary histories of a genome. *Nucleic Acids Research,* 42**,** D897-D902.

ISHIZAKI, Y., TSUNOYAMA, Y., HATANO, K., ANDO, K., KATO, K., SHINMYO, A., KOBORI, M., TAKEBA, G., NAKAHIRA, Y. & SHIINA, T. 2005. A nuclear-encoded sigma factor, Arabidopsis SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *Plant Journal,* 42**,** 133-144.

ISOM, D. G., CASTAÑEDA, C. A., CANNON, B. R. & GARCÍA-MORENO E, B. 2011. Large shifts in pK(a) values of lysine residues buried inside a protein. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 5260-5265.

JAMES, A. B., SYED, N. H., BROWN, J. W. S. & NIMMO, H. G. 2012. Thermoplasticity in the plant circadian clock. *Plant Signaling & Behavior,* 7**,** 1219-1223.

JARVIS, P. & LOPEZ-JUEZ, E. 2013. Biogenesis and homeostasis of chloroplasts and other plastids. *Nature Reviews Molecular Cell Biology,* 14**,** 787-802.

JOHANSSON, M., RAMOS-SÁNCHEZ, J. M., CONDE, D., IBÁÑEZ, C., TAKATA, N., ALLONA, I. & ERIKSSON, M. E. 2015. Role of the circadian clock in cold acclimation and winter dormancy in perennial plants. 51-74.

JOHNSON, W. C., MORAN JR, C. P. & LOSICK, R. 1983. Two RNA polymerase sigma factors from Bacillus subtilis discriminate between overlapping promoters for a developmentally regulated gene. *Nature,* 302**,** 800.

JOVANOVIC, M., ROONEY, M. S., MERTINS, P., PRZYBYLSKI, D., CHEVRIER, N., SATIJA, R., RODRIGUEZ, E. H., FIELDS, A. P., SCHWARTZ, S., RAYCHOWDHURY, R., MUMBACH, M. R., EISENHAURE, T., RABANI, M., GENNERT, D., LU, D., DELOREY, T., WEISSMAN, J. S., CARR, S. A., HACOHEN, N. & REGEV, A. 2015. Dynamic profiling of the protein life cycle in response to pathogens. *Science,* 347.

KANAMARU, K., NAGASHIMA, A., FUJIWARA, M., SHIMADA, H., SHIRANO, Y., NAKABAYASHI, K., SHIBATA, D., TANAKA, K. & TAKAHASHI, H. 2001. An Arabidopsis sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. *Plant Cell Physiology,* 42**,** 1034-43.

KANAZAWA, T., ISHIZAKI, K., KOHCHI, T., HANAOKA, M. & TANAKA, K. 2013. Characterization of four nuclear-encoded plastid RNA polymerase sigma factor genes in the liverwort *Marchantia polymorpha*: Blue-light- and multiple stress-responsive *SIG5* was acquired early in the emergence of terrestrial plants. *Plant and Cell Physiology,* 54**,** 1736-1748.

KAPLAN, F., KOPKA, J., SUNG, D. Y., ZHAO, W., POPP, M., PORAT, R. & GUY, C. L. 2007. Transcript and metabolite profiling during cold acclimation of Arabidopsis reveals an intricate relationship of cold-regulated gene expression with modifications in metabolite content. *Plant Journal,* 50**,** 967-981.

KARPINSKI, S., ESCOBAR, C., KARPINSKA, B., CREISSEN, G. & MULLINEAUX, P. M. 1997. Photosynthetic electron transport regulates the expression of cytosolic ascorbate peroxidase genes in Arabidopsis during excess light stress. *Plant Cell,* 9**,** 627.

KAWAGOE, T. & KUDOH, H. 2010. Escape from floral herbivory by early flowering in Arabidopsis halleri subsp. gemmifera. *Oecologia,* 164**,** 713-20.

KENNEDY, P. 2005. Oh no! I got the wrong sign! What should I do? *The Journal of Economic Education,* 36**,** 77-92.

KIANIANMOMENI, A. 2014. Cell-type specific light-mediated transcript regulation in the multicellular alga Volvox carteri. *BMC Genomics,* 15**,** 764.

KILIAN, J., WHITEHEAD, D., HORAK, J., WANKE, D., WEINL, S., BATISTIC, O., D'ANGELO, C., BORNBERG-BAUER, E., KUDLA, J. & HARTER, K. 2007. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant Journal,* 50**,** 347-363.

KIM, W.-Y., FUJIWARA, S., SUH, S.-S., KIM, J., KIM, Y., HAN, L., DAVID, K., PUTTERILL, J., NAM, H. G. & SOMERS, D. E. 2007. ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. *Nature,* 449**,** 356.

KIM, W.-Y., GENG, R. & SOMERS, D. E. 2003. Circadian phase-specific degradation of the F-box protein ZTL is mediated by the proteasome. *Proceedings of the National Academy of Sciences,* 100**,** 4933.

KLEINE, T. & LEISTER, D. 2016. Retrograde signaling: Organelles go networking. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1857**,** 1313-1325.

KODAMA, Y., TSUBOI, H., KAGAWA, T. & WADA, M. 2008. Low temperature-induced chloroplast relocation mediated by a blue light receptor, phototropin 2, in fern gametophytes. *Journal of Plant Research,* 121**,** 441-448.

KOLLER, B., GINGRICH, J. C., STIEGLER, G. L., FARLEY, M. A., DELIUS, H. & HALLICK, R. B. 1984. Nine introns with conserved boundary sequences in the Euglena gracilis chloroplast ribulose-1,5-bisphosphate carboxylase gene. *Cell,* 36**,** 545-553.

KOUSSEVITZKY, S., NOTT, A., MOCKLER, T. C., HONG, F., SACHETTO-MARTINS, G., SURPIN, M., LIM, J., MITTLER, R. & CHORY, J. 2007. Signals from chloroplasts converge to regulate nuclear gene expression. *Science,* 316**,** 715.

KREPS, J. A., WU, Y., CHANG, H. S., ZHU, T., WANG, X. & HARPER, J. F. 2002. Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. *Plant Physiology,* 130**,** 2129-41.

KRYNICKA, V., SHAO, S., NIXON, P. J. & KOMENDA, J. 2015. Accessibility controls selective degradation of photosystem II subunits by FtsH protease. *Nature Plants,* 1**,** 15168.

KUDOH, H. 2015. Molecular phenology in plants: in natura systems biology for the comprehensive understanding of seasonal responses under natural environments. *New Phytologist,* 210**,** 399-412.

KUDOH, H. 2016. Molecular phenology in plants: in natura systems biology for the comprehensive understanding of seasonal responses under natural environments. *New Phytologist,* 210**,** 399-412.

KUDOH, H. & NAGANO, A. J. 2013. Memory of temperature in the seasonal control of flowering time: An unexplored link between meteorology and molecular biology. *In:* PONTAROTTI, P. (ed.) *Evolutionary Biology: Exobiology and Evolutionary Mechanisms.* Berlin, Heidelberg: Springer Berlin Heidelberg.

KUREPIN, L. V., DAHAL, K. P., SAVITCH, L. V., SINGH, J., BODE, R., IVANOV, A. G., HURRY, V. & HUNER, N. P. 2013. Role of CBFs as integrators of chloroplast redox, phytochrome and plant hormone signaling during cold acclimation. *International Journal of Molecular Sciences,* 14**,** 12729-63.

KWON, Y.-J., PARK, M.-J., KIM, S.-G., BALDWIN, I. T. & PARK, C.-M. 2014. Alternative splicing and nonsense-mediated decay of circadian clock genes under environmental stress conditions in Arabidopsis. *BMC Plant Biology,* 14**,** 136.

LACKNER, D. H., SCHMIDT, M. W., WU, S., WOLF, D. A. & BÄHLER, J. 2012. Regulation of transcriptome, translation, and proteome in response to environmental stress in fission yeast. *Genome Biology,* 13**,** R25.

LANGE, B. M., RUJAN, T., MARTIN, W. & CROTEAU, R. 2000. Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proceedings of the National Academy of Sciences,* 97**,** 13172.

LI, J. J., BICKEL, P. J. & BIGGIN, M. D. 2014. System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ,* 2**,** e270.

LINDE, A. M., EKLUND, D. M., KUBOTA, A., PEDERSON ERIC, R. A., HOLM, K., GYLLENSTRAND, N., NISHIHAMA, R., CRONBERG, N., MURANAKA, T., OYAMA, T., KOHCHI, T. & LAGERCRANTZ, U. 2017. Early evolution of the land plant circadian clock. *New Phytologist,* 216**,** 576-590.

LITTHAUER, S., BATTLE MARTIN, W., LAWSON, T. & JONES MATTHEW, A. 2015. Phototropins maintain robust circadian oscillation of PSII operating efficiency under blue light. *Plant Journal,* 83**,** 1034-1045.

LITTHAUER, S., CHAN, K. X. & JONES, M. A. 2018. 3′-Phosphoadenosine 5′-Phosphate accumulation delays the circadian system. *Plant Physiology,* 176**,** 3120-3135.

LIU, B., YANG, Z., GOMEZ, A., LIU, B., LIN, C. & OKA, Y. 2016a. Signaling mechanisms of plant cryptochromes in *Arabidopsis thaliana*. *Journal of Plant Research,* 129**,** 137-148.

LIU, H., YU, X., LI, K., KLEJNOT, J., YANG, H., LISIERO, D. & LIN, C. 2008. Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. *Science,* 322**,** 1535.

LIU, Y., BEYER, A. & AEBERSOLD, R. 2016b. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell,* 165**,** 535-550.

LIVINGSTONE, C. D. & BARTON, G. J. 1993. Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *Bioinformatics,* 9**,** 745-756.

MALAPEIRA, J., KHAITOVA, L. C. & MAS, P. 2012. Ordered changes in histone modifications at the core of the Arabidopsis circadian clock. *Proceedings of the National Academy of Sciences,* 109**,** 21540-21545.

MARCOVITZ, A. & LEVY, Y. 2011. Frustration in protein–DNA binding influences conformational switching and target search kinetics. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 17957-17962.

MAXWELL, K. & JOHNSON, G. N. 2000. Chlorophyll fluorescence—a practical guide. *Journal of Experimental Botany,* 51**,** 659-668.

MAY, T. & SOLL, J. 2000. 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. *Plant Cell,* 12**,** 53-63.

MCCLUNG, C. R. 2006. Plant Circadian Rhythms. *Plant Cell,* 18**,** 792-803.

MCCLUNG, C. R., SALOMÉ, P. A. & MICHAEL, T. P. 2002. The Arabidopsis Circadian System. *The Arabidopsis Book / American Society of Plant Biologists,* 1**,** e0044.

MCCORMAC, A. C. & TERRY, M. J. 2004. The nuclear genes Lhcb and HEMA1 are differentially sensitive to plastid signals and suggest distinct roles for the GUN1 and GUN5 plastid-signalling pathways during de-etiolation. *Plant Journal,* 40**,** 672-685.

MCFADDEN, G. I. & VAN DOOREN, G. G. 2004. Evolution: Red Algal Genome Affirms a Common Origin of All Plastids. *Current Biology,* 14**,** R514-R516.

MCWATTERS, H. G. & DEVLIN, P. F. 2011. Timing in plants – A rhythmic arrangement. *FEBS Letters,* 585**,** 1474-1484.

MERESCHKOWSKY, C. 1920. La plante considére comme un complexe symbiotique. *Bulletin de la Société des sciences naturelles de l'Ouest de la France,* 6**,** 17-98.

MESKAUSKIENE, R., NATER, M., GOSLINGS, D., KESSLER, F., OP DEN CAMP, R. & APEL, K. 2001. FLU: a negative regulator of chlorophyll biosynthesis in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences,* 98**,** 12826.

MICHAELS, S. D. & AMASINO, R. M. 1999. FLOWERING LOCUS C Encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell,* 11**,** 949.

MILLAR, A. J., CARRE, I. A., STRAYER, C. A., CHUA, N. H. & KAY, S. A. 1995. Circadian clock mutants in Arabidopsis identified by luciferase imaging. *Science,* 267**,** 1161.

MITSUI, A., KUMAZAWA, S., TAKAHASHI, A., IKEMOTO, H., CAO, S. & ARAI, T. 1986. Strategy by which nitrogen-fixing unicellular cyanobacteria grow photoautotrophically. *Nature,* 323**,** 720.

MIZUNO, T., NOMOTO, Y., OKA, H., KITAYAMA, M., TAKEUCHI, A., TSUBOUCHI, M. & YAMASHINO, T. 2014. Ambient temperature signal feeds into the circadian clock transcriptional circuitry through the EC night-time repressor in Arabidopsis thaliana. *Plant Cell Physiology,* 55**,** 958-76.

MOCHIZUKI, N., BRUSSLAN, J. A., LARKIN, R., NAGATANI, A. & CHORY, J. 2001. Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proceedings of the National Academy of Sciences,* 98**,** 2053.

MOCHIZUKI, N., TANAKA, R., TANAKA, A., MASUDA, T. & NAGATANI, A. 2008. The steady-state level of Mg-protoporphyrin IX is not a determinant of plastid-to-nucleus signaling in Arabidopsis. *Proceedings of the National Academy of Sciences,* 105**,** 15184.

MOCKLER, T. C., MICHAEL, T. P., PRIEST, H. D., SHEN, R., SULLIVAN, C. M., GIVAN, S. A., MCENTEE, C., KAY, S. A. & CHORY, J. 2007. The diurnal project: Diurnal and circadian expression profiling, model-based pattern matching, and promoter analysis. *Cold Spring Harbor Symposia on Quantitative Biology,* 72**,** 353-363.

MOELLERING, E. R., MUTHAN, B. & BENNING, C. 2010. Freezing tolerance in plants requires lipid remodeling at the outer chloroplast membrane. *Science,* 330**,** 226.

MOLINA, J., HAZZOURI, K. M., NICKRENT, D., GEISLER, M., MEYER, R. S., PENTONY, M. M., FLOWERS, J. M., PELSER, P., BARCELONA, J., INOVEJAS, S. A., UY, I., YUAN, W., WILKINS, O., MICHEL, C.-I., LOCKLEAR, S., CONCEPCION, G. P. & PURUGGANAN, M. D. 2014. Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Molecular Biology and Evolution,* 31**,** 793-803.

MONGRAND, S., BESSOULE, J.-J., CABANTOUS, F. & CASSAGNE, C. 1998. The C16:3/C18:3 fatty acid balance in photosynthetic tissues from 468 plant species. *Phytochemistry,* 49**,** 1049-1064.

MOON, B. Y., HIGASHI, S., GOMBOS, Z. & MURATA, N. 1995. Unsaturation of the membrane lipids of chloroplasts stabilizes the photosynthetic machinery against low-temperature photoinhibition in transgenic tobacco plants. *Proceedings of the National Academy of Sciences,* 92**,** 6219-6223.

MORONEY, J. V., BARTLETT, S. G. & SAMUELSSON, G. 2001. Carbonic anhydrases in plants and algae: Invited review. *Plant, Cell & Environment,* 24**,** 141-153.

MURANAKA, T. & OYAMA, T. 2016. Heterogeneity of cellular circadian clocks in intact plants and its correction under light-dark cycles. *Science Advances,* 2.

MURATA, N., TAKAHASHI, S., NISHIYAMA, Y. & ALLAKHVERDIEV, S. I. 2007. Photoinhibition of photosystem II under environmental stress. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1767**,** 414-21.

NAGANO, A. J., SATO, Y., MIHARA, M., ANTONIO, B. A., MOTOYAMA, R., ITOH, H., NAGAMURA, Y. & IZAWA, T. 2012. Deciphering and prediction of transcriptome dynamics under fluctuating field conditions. *Cell,* 151**,** 1358-69.

NAGASHIMA, A., HANAOKA, M., SHIKANAI, T., FUJIWARA, M., KANAMARU, K., TAKAHASHI, H. & TANAKA, K. 2004. The multiple-stress responsive plastid sigma factor, SIG5, directs activation of the psbD Blue Light-Responsive Promoter (BLRP) in *Arabidopsis thaliana*. *Plant and Cell Physiology,* 45**,** 357-368.

NAGEL, D. H. & KAY, S. A. 2012. Complexity in the wiring and regulation of plant circadian networks. *Current Biology,* 22**,** R648-57.

NAKAMICHI, N., KIBA, T., HENRIQUES, R., MIZUNO, T., CHUA, N.-H. & SAKAKIBARA, H. 2010. PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. *Plant Cell,* 22**,** 594-605.

NOORDALLY, Z. B., ISHII, K., ATKINS, K. A., WETHERILL, S. J., KUSAKINA, J., WALTON, E. J., KATO, M., AZUMA, M., TANAKA, K., HANAOKA, M. & DODD, A. N. 2013. Circadian control of chloroplast transcription by a nuclear-encoded timing signal. *Science,* 339**,** 1316-9.

NOORDALLY, Z. B. & MILLAR, A. J. 2015. Clocks in algae. *Biochemistry,* 54**,** 171-83.

OELMÜLLER, R., LEVITAN, I., BERGFELD, R., RAJASEKHAR, V. K. & MOHR, H. 1986. Expression of nuclear genes as affected by treatments acting on the plastids. *Planta,* 168**,** 482-492.

OELMÜLLER, R. & MOHR, H. 1986. Photooxidative destruction of chloroplasts and its consequences for expression of nuclear genes. *Planta,* 167**,** 106-113.

OHARA, T., FUKUDA, H. & TOKUDA, I. T. 2015. Phase response of the *Arabidopsis thaliana* circadian clock to light pulses of different wavelengths. *Journal of Biological Rhythms,* 30**,** 95-103.

OKADA, M., INOUE, M. & IKEDA, T. 1978. Circadian rhythm in photosynthesis of the green alga Bryopsis maxima. *Plant and Cell Physiology,* 19**,** 197-202.

OLSSON, M. H. M., SØNDERGAARD, C. R., ROSTKOWSKI, M. & JENSEN, J. H. 2011. PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *Journal of Chemical Theory and Computation,* 7**,** 525-537.

ONDA, Y., YAGI, Y., SAITO, Y., TAKENAKA, N. & TOYOSHIMA, Y. 2008. Light induction of Arabidopsis SIG1 and SIG5 transcripts in mature leaves: differential roles of cryptochrome 1 and cryptochrome 2 and dual function of SIG5 in the recognition of plastid promoters. *Plant Journal,* 55**,** 968-978.

OP DEN CAMP, R. G. L., PRZYBYLA, D., OCHSENBEIN, C., LALOI, C., KIM, C., DANON, A., WAGNER, D., HIDEG, É., GÖBEL, C., FEUSSNER, I., NATER, M. & APEL, K. 2003. Rapid induction of distinct stress responses after the release of singlet oxygen in Arabidopsis. *Plant Cell,* 15**,** 2320.

PAGE, M. T., KACPRZAK, S. M., MOCHIZUKI, N., OKAMOTO, H., SMITH, A. G. & TERRY, M. J. 2017. Seedlings lacking the PTM protein do not show a genomes uncoupled (*gun*) mutant phenotype. *Plant Physiology,* 174**,** 21.

PAGET, M. S. & HELMANN, J. D. 2003. The σ 70 family of sigma factors. *Genome biology,* 4**,** 203.

PANDA, S., ANTOCH, M. P., MILLER, B. H., SU, A. I., SCHOOK, A. B., STRAUME, M., SCHULTZ, P. G., KAY, S. A., TAKAHASHI, J. S. & HOGENESCH, J. B. 2002a. Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell,* 109**,** 307-320.

PANDA, S., HOGENESCH, J. B. & KAY, S. A. 2002b. Circadian rhythms from flies to human. *Nature,* 417**,** 329.

PENFIELD, S. 2008. Temperature perception and signal transduction in plants. *New Phytologist,* 179**,** 615-28.

PLAUT, W. & SAGAN, L. A. 1958. Incorporation of thymidine in the cytoplasm of *Amoeba proteus*. *The Journal of Biophysical and Biochemical Cytology,* 4**,** 843.

PRUNEDA-PAZ, J. L., BRETON, G., PARA, A. & KAY, S. A. 2009. A Functional Genomics Approach Reveals CHE as a Component of the Arabidopsis Circadian Clock. *Science,* 323**,** 1481-1485.

R CORE TEAM 2014. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

RAMEL, F., BIRTIC, S., GINIES, C., SOUBIGOU-TACONNAT, L., TRIANTAPHYLIDÈS, C. & HAVAUX, M. 2012. Carotenoid oxidation products are stress signals that mediate gene responses to singlet oxygen in plants. *Proceedings of the National Academy of Sciences,* 109**,** 5535.

RAWAT, R., TAKAHASHI, N., HSU, P. Y., JONES, M. A., SCHWARTZ, J., SALEMI, M. R., PHINNEY, B. S. & HARMER, S. L. 2011. REVEILLE8 and PSEUDO-REPONSE REGULATOR5 form a negative feedback loop within the Arabidopsis circadian clock. *PLOS Genetics,* 7**,** e1001350.

REYES-PRIETO, A., WEBER, A. P. M. & BHATTACHARYA, D. 2007. The origin and establishment of the plastid in algae and plants. *Annual Review of Genetics,* 41**,** 147-168.

RIZZINI, L., FAVORY, J.-J., CLOIX, C., FAGGIONATO, D., O'HARA, A., KAISERLI, E., BAUMEISTER, R., SCHÄFER, E., NAGY, F., JENKINS, G. I. & ULM, R. 2011. Perception of UV-B by the Arabidopsis UVR8 protein. *Science,* 332**,** 103.

ROBLES, M. S., COX, J. & MANN, M. 2014. In-Vivo Quantitative Proteomics Reveals a Key Contribution of Post-Transcriptional Mechanisms to the Circadian Regulation of Liver Metabolism. *PLOS Genetics,* 10**,** e1004047.

ROGERS, M. B., GILSON, P. R., SU, V., MCFADDEN, G. I. & KEELING, P. J. 2007. The complete chloroplast genome of the Chlorarachniophyte *Bigelowiella natans*: Evidence for independent origins of Chlorarachniophyte and Euglenid secondary endosymbionts. *Molecular Biology and Evolution,* 24**,** 54-62.

ROHS, R., JIN, X., WEST, S. M., JOSHI, R., HONIG, B. & MANN, R. S. 2010. Origins of specificity in protein-DNA recognition. *Annual Review of Biochemistry,* 79**,** 233-269.

ROSSEL, J. B., WALTER, P. B., HENDRICKSON, L., CHOW, W. S., POOLE, A., MULLINEAUX, P. M. & POGSON, B. J. 2005. A mutation affecting ASCORBATE PEROXIDASE 2 gene expression reveals a link between responses to high light and drought tolerance. *Plant, Cell & Environment,* 29**,** 269-281.

SALOME, P. A. & MCCLUNG, C. R. 2005. PSEUDO-RESPONSE REGULATOR 7 and 9 are partially redundant genes essential for the temperature responsiveness of the Arabidopsis circadian clock. *Plant Cell,* 17**,** 791-803.

SANCHEZ, S. E., PETRILLO, E., BECKWITH, E. J., ZHANG, X., RUGNONE, M. L., HERNANDO, C. E., CUEVAS, J. C., GODOY HERZ, M. A., DEPETRIS-CHAUVIN, A., SIMPSON, C. G., BROWN, J. W. S., CERDAN, P. D., BOREVITZ, J. O., MAS, P., CERIANI, M. F., KORNBLIHTT, A. R. & YANOVSKY, M. J. 2010. A methyl transferase links the circadian clock to the regulation of alternative splicing. *Nature,* 468**,** 112-116.

SATAKE, A., KAWAGOE, T., SABURI, Y., CHIBA, Y., SAKURAI, G. & KUDOH, H. 2013. Forecasting flowering phenology under climate warming by modelling the regulatory dynamics of flowering-time genes. *Nature Communications,* 4**,** 2303.

SCHMID, R. & DRING, M. J. 1996. Influence of carbon supply on the circadian rhythmicity of photosynthesis and its stimulation by blue light in Ectocarpus siliculosus: clues to the mechanism of inorganic carbon acquisition in lower brown algae. *Plant, Cell & Environment,* 19**,** 373-382.

SCHMID, R., DRING, M. J. & FORSTER, R. M. 1994. Kinetics of blue-light stimulation and circadian rhythmicity of light-saturated photosynthesis in brown algae: a species comparison. *Journal of Phycology,* 30**,** 612-621.

SCHMITZ-LINNEWEBER, C., WILLIAMS-CARRIER, R. E., WILLIAMS-VOELKER, P. M., KROEGER, T. S., VICHAS, A. & BARKAN, A. 2006. A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 pre-mRNA. *Plant Cell,* 18**,** 2650.

SCHRODINGER, LLC 2015. The PyMOL Molecular Graphics System, Version 1.8.

SCHWEER, J., TÜRKERI, H., LINK, B. & LINK, G. 2010. AtSIG6, a plastid sigma factor from Arabidopsis, reveals functional impact of cpCK2 phosphorylation. *Plant Journal,* 62**,** 192-202.

SEARLE, I., HE, Y., TURCK, F., VINCENT, C., FORNARA, F., KRÖBER, S., AMASINO, R. A. & COUPLAND, G. 2006. The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. *Genes & Development,* 20**,** 898-912.

SEATON, D. D., SMITH, R. W., SONG, Y. H., MACGREGOR, D. R., STEWART, K., STEEL, G., FOREMAN, J., PENFIELD, S., IMAIZUMI, T., MILLAR, A. J. & HALLIDAY, K. J. 2015. Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature. *Mol Syst Biol,* 11**,** 776.

SENN, I. G. 1909. Weitere Untersuchungen über die Gestaltsund Lageveränderung der Chromatophoren. *Berichte der Deutschen Botanischen Gesellschaft,* 27**,** 12-27.

SEO, P. J. & MAS, P. 2015. STRESSing the role of the plant circadian clock. *Trends in Plant Science,* 20**,** 230-7.

SEO, P. J., PARK, M.-J., LIM, M.-H., KIM, S.-G., LEE, M., BALDWIN, I. T. & PARK, C.-M. 2012. A self-regulatory circuit of CIRCADIAN CLOCK-ASSOCIATED1 underlies the circadian clock regulation of temperature responses in Arabidopsis. *Plant Cell*.

SERRANO-BUENO, G., ROMERO-CAMPERO, F. J., LUCAS-REINA, E., ROMERO, J. M. & VALVERDE, F. 2017. Evolution of photoperiod sensing in plants and algae. *Current Opinion in Plant Biology,* 37**,** 10-17.

SHIKANAI, T. 2016. Chloroplast NDH: A different enzyme with a structure similar to that of respiratory NADH dehydrogenase. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 1857**,** 1015-1022.

SHIMIZU, M., KATO, H., OGAWA, T., KURACHI, A., NAKAGAWA, Y. & KOBAYASHI, H. 2010. Sigma factor phosphorylation in the photosynthetic control of photosystem stoichiometry. *Proceedings of the National Academy of Sciences,* 107**,** 10760-10764.

SOEJIMA, A., YAMAZAKI, N., NISHINO, T. & WAKANA, I. 2008. Genetic variation and structure of the endangered freshwater benthic alga Marimo, *Aegagropila linnaei* (Ulvophyceae) in Japanese lakes. *Aquatic Ecology,* 43**,** 359-370.

STEPONKUS, P. L. 1984. Role of the plasma membrane in freezing injury and cold acclimation. *Annual Review of Plant Physiology,* 35**,** 543-584.

STEPONKUS, P. L. & LYNCH, D. V. 1989a. Freeze/thaw-induced destabilization of the plasma membrane and the effects of cold acclimation. *Journal of Bioenergetics and Biomembranes,* 21**,** 21-41.

STEPONKUS, P. L. & LYNCH, D. V. 1989b. Freeze/thaw-induced destabilization of the plasma membrane and the effects of cold acclimation. *Journal of Bioenergetics and Biomembranes,* 21**,** 21-41.

STEPONKUS, P. L., UEMURA, M., JOSEPH, R. A., GILMOUR, S. J. & THOMASHOW, M. F. 1998a. Mode of action of the COR15a gene on the freezing tolerance of Arabidopsis thaliana. *Proceedings of the National Academy of Sciences,* 95**,** 14570-14575.

STEPONKUS, P. L., UEMURA, M., JOSEPH, R. A., GILMOUR, S. J. & THOMASHOW, M. F. 1998b. Mode of action of the COR15a gene on the freezing tolerance of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 14570-14575.

STRAND, Å., ASAMI, T., ALONSO, J., ECKER, J. R. & CHORY, J. 2003. Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. *Nature,* 421**,** 79.

SUAREZ-LOPEZ, P., WHEATLEY, K., ROBSON, F., ONOUCHI, H., VALVERDE, F. & COUPLAND, G. 2001. CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature,* 410**,** 1116-1120.

SUN, X., FENG, P., XU, X., GUO, H., MA, J., CHI, W., LIN, R., LU, C. & ZHANG, L. 2011. A chloroplast envelope-bound PHD transcription factor mediates chloroplast signals to the nucleus. *Nature Communications,* 2**,** 477.

SUSEK, R. E., AUSUBEL, F. M. & CHORY, J. 1993. Signal transduction mutants of arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell,* 74**,** 787-799.

SWEENEY, B. M., TUFFLI, C. F. & RUBIN, R. H. 1967. The circadian rhythm in photosynthesis in Acetabularia in the presence of actinomycin D, puromycin, and chloramphenicol. *The Journal of General Physiology,* 50**,** 647-659.

TADINI, L., PESARESI, P., KLEINE, T., ROSSI, F., GULJAMOW, A., SOMMER, F., MÜHLHAUS, T., SCHRODA, M., MASIERO, S., PRIBIL, M., ROTHBART, M., HEDTKE, B., GRIMM, B. & LEISTER, D. 2016. GUN1 controls accumulation of the plastid ribosomal protein S1 at the protein level and interacts with proteins involved in plastid protein homeostasis. *Plant Physiology,* 170**,** 1817.

TAKAHASHI, N., HIRATA, Y., AIHARA, K. & MAS, P. 2015. A hierarchical multi-oscillator network orchestrates the Arabidopsis circadian system. *Cell,* 163**,** 148-159.

TAKAHASHI, S. & MURATA, N. 2008. How do environmental stresses accelerate photoinhibition? *Trends in Plant Science,* 13**,** 178-182.

TAKEUCHI, T., NEWTON, L., BURKHARDT, A., MASON, S. & FARRÉ, E. M. 2014. Light and the circadian clock mediate time-specific changes in sensitivity to UV-B stress under light/dark cycles. *Journal of Experimental Botany,* 65**,** 6003-6012.

TALARICO, L. & MARANZANA, G. 2000. Light and adaptive responses in red macroalgae: an overview. *Journal of Photochemistry and Photobiology B: Biology,* 56**,** 1-11.

TANAKA, K., TOZAWA, Y., MOCHIZUKI, N., SHINOZAKI, K., NAGATANI, A., WAKASA, K. & TAKAHASHI, H. 1997. Characterization of three cDNA species encoding plastid RNA polymerase sigma factors in Arabidopsis thaliana: Evidence for the sigma factor heterogeneity in higher plant plastids. *FEBS Letters,* 413**,** 309-313.

THOMASHOW, M. F. 1999. Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology,* 50**,** 571-599.

THOMASHOW, M. F. 2010. Molecular basis of plant cold acclimation: insights gained from studying the CBF cold response pathway. *Plant Physiology,* 154**,** 571-7.

THORLBY, G., FOURRIER, N. & WARREN, G. 2004. The Sensitive to Freezing2 gene, required for freezing tolerance in a *Arabidopsis thaliana*, encodes a β-glucosidase. *Plant Cell,* 16**,** 2192-2203.

TOGASHI, T., SASAKI, H. & YOSHIMURA, J. 2014. A geometrical approach explains Lake Ball (Marimo) formations in the green alga, Aegagropila linnaei. *Scientific Reports,* 4**,** 3761.

TORABI, S., UMATE, P., MANAVSKI, N., PLÖCHINGER, M., KLEINKNECHT, L., BOGIREDDI, H., HERRMANN, R. G., WANNER, G., SCHRÖDER, W. P. & MEURER, J. 2014. PsbN is required for assembly of the photosystem II reaction center in *Nicotiana tabacum*. *Plant Cell,* 26**,** 1183.

TSUNOYAMA, Y., ISHIZAKI, Y., MORIKAWA, K., KOBORI, M., NAKAHIRA, Y., TAKEBA, G., TOYOSHIMA, Y. & SHIINA, T. 2004. Blue light-induced transcription of plastid-encoded psbD gene is mediated by a nuclear-encoded transcription initiation factor, AtSig5. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 3304-9.

VADEBONCOEUR, Y., PETERSON, G., VANDER ZANDEN, M. J. & KALFF, J. 2008. Benthic algal production across lake size gradients: interactions among morphometry, nutrients, and light. *Ecology,* 89**,** 2542-2552.

WAEGEMANN, K. & SOLL, J. 1996. Phosphorylation of the transit sequence of chloroplast precursor proteins. *Journal of Biological Chemistry,* 271**,** 6545-6554.

WAGENMAKERS, E.-J. & FARRELL, S. 2004. AIC model selection using Akaike weights. *Psychonomic Bulletin & Review,* 11**,** 192-196.

WAGNER, D., PRZYBYLA, D., OP DEN CAMP, R., KIM, C., LANDGRAF, F., LEE, K. P., WÜRSCH, M., LALOI, C., NATER, M., HIDEG, E. & APEL, K. 2004. The genetic basis of singlet oxygen-induced stress responses of *Arabidopsis thaliana*. *Science,* 306**,** 1183.

WANG, H., MA, L.-G., LI, J.-M., ZHAO, H.-Y. & DENG, X. W. 2001. Direct Interaction of Arabidopsis Cryptochromes with COP1 in Light Control Development. *Science,* 294**,** 154.

WATERHOUSE, A. M., PROCTER, J. B., MARTIN, D. M. A., CLAMP, M. & BARTON, G. J. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* 25**,** 1189-1191.

WATERS, M. T., WANG, P., KORKARIC, M., CAPPER, R. G., SAUNDERS, N. J. & LANGDALE, J. A. 2009. GLK Transcription Factors Coordinate Expression of the Photosynthetic Apparatus in <em>Arabidopsis</em&gt. *Plant Cell,* 21**,** 1109.

WEIHE, A. & BÖRNER, T. 1999. Transcription and the architecture of promoters in chloroplasts. *Trends in Plant Science,* 4**,** 169-170.

WENDEN, B., KOZMA-BOGNAR, L., EDWARDS, K. D., HALL, A. J., LOCKE, J. C. & MILLAR, A. J. 2011. Light inputs shape the Arabidopsis circadian system. *Plant Journal,* 66**,** 480-91.

WILSON, P. B., ESTAVILLO, G. M., FIELD, K. J., PORNSIRIWONG, W., CARROLL, A. J., HOWELL, K. A., WOO, N. S., LAKE, J. A., SMITH, S. M., HARVEY MILLAR, A., VON CAEMMERER, S. & POGSON, B. J. 2009. The nucleotidase/phosphatase SAL1 is a negative regulator of drought tolerance in Arabidopsis. *Plant Journal,* 58**,** 299-317.

WINTER, D., VINEGAR, B., NAHAL, H., AMMAR, R., WILSON, G. V. & PROVART, N. J. 2007a. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLOS One,* 2**,** e718.

WINTER, D., VINEGAR, B., NAHAL, H., AMMAR, R., WILSON, G. V. & PROVART, N. J. 2007b. An “Electronic Fluorescent Pictograph” Browser for Exploring and Analyzing Large-Scale Biological Data Sets. *PLOS One,* 2**,** e718.

WOODSON, JESSE D., PEREZ-RUIZ, JUAN M. & CHORY, J. 2011. Heme Synthesis by Plastid Ferrochelatase I Regulates Nuclear Gene Expression in Plants. *Current Biology,* 21**,** 897-903.

WOODSON, J. D., PEREZ-RUIZ, J. M., SCHMITZ, R. J., ECKER, J. R. & CHORY, J. 2013. Sigma factor-mediated plastid retrograde signals control nuclear gene expression. *Plant Journal,* 73**,** 1-13.

XIAO, Y., SAVCHENKO, T., BAIDOO, EDWARD E. K., CHEHAB, WASSIM E., HAYDEN, DANIEL M., TOLSTIKOV, V., CORWIN, JASON A., KLIEBENSTEIN, DANIEL J., KEASLING, JAY D. & DEHESH, K. 2012. Retrograde Signaling by the

Plastidial Metabolite MEcPP Regulates Expression of Nuclear Stress-Response Genes. *Cell,* 149**,** 1525-1535.

XU, M. Q., KATHE, S. D., GOODRICH-BLAIR, H., NIERZWICKI-BAUER, S. A. & SHUB, D. A. 1990. Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacteria. *Science,* 250**,** 1566.

YANG, J., YAN, R., ROY, A., XU, D., POISSON, J. & ZHANG, Y. 2014. The I-TASSER Suite: protein structure and function prediction. *Nature Methods,* 12**,** 7.

YANOVSKY, M. J. & KAY, S. A. 2002. Molecular basis of seasonal time measurement in Arabidopsis. *Nature,* 419**,** 308.

YESUDHAS, D., BATOOL, M., ANWAR, M. A., PANNEERSELVAM, S. & CHOI, S. 2017. Proteins recognizing DNA: Structural uniqueness and versatility of DNA-binding domains in stem cell transcription factors. *Genes,* 8**,** 192.

ZGHIDI, W., MERENDINO, L., COTTET, A., MACHE, R. & LERBS-MACHE, S. 2007. Nucleus-encoded plastid sigma factor SIG3 transcribes specifically the psbN gene in plastids. *Nucleic Acids Research,* 35**,** 455-64.

ZHANG, Y., XI, Z., HEGDE, R. S., SHAKKED, Z. & CROTHERS, D. M. 2004. Predicting indirect readout effects in protein–DNA interactions. *Proceedings of the National Academy of Sciences,* 101**,** 8337-8341.

ZHENG, G., LI, L. & LI, W. 2016. Glycerolipidome responses to freezing- and chilling-induced injuries: examples in Arabidopsis and rice. *BMC Plant Biology,* 16**,** 70.

ZIELINSKI, T., MOORE, A. M., TROUP, E., HALLIDAY, K. J. & MILLAR, A. J. 2014. Strengths and limitations of period estimation methods for circadian data. *PLOS One,* 9**,** e96462.